

Алгоритмы построения многокадрового сверхразрешения изображений в условиях аппликативных помех на основе глубоких нейронных сетей

С.В. Саввин¹, А.А. Сирота¹

¹ Воронежский государственный университет,
394018, Россия, г. Воронеж, Университетская площадь, д. 1

Аннотация

Рассматриваются алгоритмы построения многокадрового сверхразрешения, позволяющие восстанавливать изображения с высоким разрешением за счет накопления последовательности изображений с низким разрешением в условиях аппликативных помех. Воздействие аппликативных помех проявляется в появлении локальных участков аномальных наблюдений на каждом изображении и также является фактором понижения разрешения. Решению данной задачи до настоящего времени уделялось недостаточно внимания, при этом перспективным подходом для обработки изображений, включая построение многокадрового сверхразрешения, является использование глубоких нейронных сетей. В работе рассмотрены существующие подходы к решению данной задачи и предложен новый подход, основанный на использовании нескольких сверточных нейронных сетей. Особенностью рассматриваемого подхода и реализуемых на его основе алгоритмов является выполнение итеративной обработки входной последовательности изображений с низким разрешением с применением нейронных сетей на разных этапах обработки, включая регистрацию изображений низкого разрешения, сегментацию и выявление участков, пораженных аппликативными помехами, а также преобразования, направленные непосредственно на повышение разрешения. Данный подход позволяет комбинировать сильные стороны существующих аналогов и устранить их основные недостатки, связанные с необходимостью использования приближенных математических моделей данных, которые требуются для синтеза алгоритмов обработки изображений в рамках статистической теории решений. Для обновления текущей оценки изображения высокого разрешения предложена специальная сверточная нейронная сеть, организованная в виде направленного ациклического графа. Проведены экспериментальные исследования, показавшие работоспособность предложенного алгоритма и его преимущество по точности восстановления изображения с высоким разрешением по сравнению с альтернативными вариантами решения задачи.

Ключевые слова: обработка изображений, сверхразрешение, сверточные нейронные сети, глубокое обучение, аппликативные помехи.

Цитирование: Саввин, С.В. Алгоритмы построения многокадрового сверхразрешения изображений в условиях аппликативных помех на основе глубоких нейронных сетей / С.В. Саввин, А.А. Сирота // Компьютерная оптика. – 2022. – Т. 46, № 1. – С. 130-138. – DOI: 10.18287/2412-6179-CO-904.

Citation: Savvin SV, Sirota AA. Algorithms for multi-frame image super-resolution under applicative noise based on deep neural networks. Computer Optics 2022; 46(1): 130-138. DOI: 10.18287/2412-6179-CO-904.

Введение

Разрешение цифровых изображений зависит от характеристик систем, формирующих и регистрирующих эти изображения. Кроме того, на разрешение изображений влияют ограничения, возникающие при их передаче по каналам связи. Для эффективной работы большинства систем обработки и анализа информации требуются изображения с высоким разрешением (ВР), которые обеспечивают требуемый уровень детализации сцен, но по ряду причин не могут быть получены аппаратно из-за ограниченных возможностей средств регистрации и передачи данных. Получить требуемый уровень детализации в этих условиях можно на основе методов сверхразрешения

(СР), улучшающих качество изображений и уровень их детализации путем их специальной обработки. В этом плане известны алгоритмы многокадрового сверхразрешения [1–15], которые позволяют получить изображение высокого разрешения (ВР) за счёт накопления последовательности изображений низкого разрешения (НР), отображающих одну и ту же сцену с субпиксельными смещениями между соседними кадрами.

Ещё одним важным фактором, определяющим качество изображений, является воздействие аппликативных помех (АП), проявляющихся в распределённых участках аномальных наблюдений, называемых локальными областями закрытия (ЛОЗ), что также можно рассматривать как фактор снижения разреше-

ния. В отличие от аддитивных помех, АП не могут быть устранены путём фильтрации изображений. В настоящее время существует множество различных подходов к многокадровому СР, однако лишь некоторые алгоритмы пытаются решить задачу повышения разрешения в условиях аппликативных искажений [8–13]. В частности, этот вопрос рассматривался в предшествующих работах авторов [13, 14], в которых для решения указанной задачи использованы итеративные алгоритмы оптимальной марковской линейной и нелинейной фильтрации. В то же время синтез подобных алгоритмов требует использования существенных допущений и приближений относительно используемых математических моделей данных и их параметров, что не всегда оправдано на практике. Например, это касается использования Гауссовской модели данных, которая в большинстве случаев не является адекватной.

В последние годы для решения задач обработки изображений всё больший интерес возникает к использованию методов глубокого машинного обучения и, в частности, свёрточных нейронных сетей и их многочисленных модификаций. В том числе глубокие нейронные сети могут успешно использоваться для решения задач многокадрового СР [3–9, 15]. Данная парадигма решения задач в области обработки изображений свободна от использования математических моделей данных, но требует существенных по объёму обучающих выборок, содержащих примеры анализируемых изображений.

Новизна предлагаемого в данной работе подхода состоит в применении методов машинного обучения, в рамках которого на всех этапах обработки используются глубокие нейронные сети, включая и этап сегментации АП на входных кадрах НР. В известных авторам работах такой подход к итеративному решению задачи построения многокадрового СР с одновременной компенсацией негативного воздействия АП на входные изображения не рассматривался. Новизной также отличается предложенная искусственная нейронная сеть с архитектурой, организованной в виде направленного ациклического графа (DAG-сеть), обеспечивающая формирование оценок ВР с повышенным качеством, а также результаты сравнения комплекса предложенных алгоритмов обработки информации с альтернативными вариантами [3, 10, 12, 13, 15].

1. Анализ предшествующих работ

Задача построения многокадрового СР находит своё отражение во множестве работ, основывающихся на различных методах и моделях: различные виды интерполяции, машинное обучение, вероятностные модели, фильтрация по Калману и т.д. Более детальные обзоры различных подходов приведены в [1, 2].

В последние годы для решения этой задачи всё чаще применяются методы, основанные на использо-

вании глубоких нейронных сетей. При этом можно выделить ряд подходов к организации обработки поступающей на вход последовательности изображений НР, которые предполагают использование нескольких нейронных сетей, обучаемых как совместно, так и отдельно друг от друга. Например, в [3] описывается итеративный алгоритм, в котором на вход глубокой свёрточной нейронной сети подаются очередной кадр НР и текущая оценка изображения ВР, а на выходе получается обновлённая оценка ВР. Регистрация входных изображений тоже осуществляется при помощи нейронной сети. Подход, представленный в [4], предполагает независимое повышение разрешения каждого входного изображения при помощи нейронной сети. После обработки всех кадров полученные изображения ВР объединяются в одно выходное изображение ВР. В [5] процесс построения СР основывается на объединении входных изображений в одно изображение ВР, детализация которого повышается при помощи свёрточной нейронной сети. Алгоритм, рассмотренный в [6], делает акцент на обработке последовательности видеок кадров (т.е. количество выходных изображений ВР равно количеству входных изображений НР). Нейронная сеть, используемая в данной работе, принимает на вход сразу несколько изображений НР: текущий кадр, а также некоторое количество кадров до и после него. В работе [7] используется схема, основанная на независимой начальной обработке каждого из входных изображений. Полученные признаки затем «сливаются» в единое остаточное изображение, которое складывается с начальной усреднённой оценкой ВР для получения выходного изображения ВР. Различные этапы обработки реализуются при помощи глубоких свёрточных нейронных сетей. Схожий подход рассматривается в [8], однако обработка полученных признаков осуществляется рекурсивным образом при помощи особым образом построенной свёрточной нейронной сети. В [9] предлагается алгоритм, основанный на парном рекурсивном слиянии изображений НР с целью получения признаков, которые затем декодируются в одно изображение ВР.

Несмотря на существование различных подходов к обработке последовательности изображений НР при помощи нейронных сетей с целью получения изображения ВР, приведённые выше алгоритмы не учитывают особенности решения данной задачи в условиях воздействия АП. Только в работе [9] поднимается вопрос о целесообразности включения в процесс построения СР результатов сегментации ложных наблюдений, однако синтезированный алгоритм не реализует в себе такой этап обработки.

Тем не менее задача построения многокадрового СР при наличии ложных наблюдений (АП) на обрабатываемых изображениях рассматривалась в работах, основанных на использовании вероятностных моделей и алгоритмов фильтрации калмановского

типа. Так, в [10] рассматривается алгоритм, который основывается на широко используемой в физике модели спиновых стёкол для моделирования АП, действующих на изображения. В [11] приводится модификация такого подхода с учётом инерционного движения АП между кадрами НР. В указанных работах для восстановления изображения ВР используется метод вариационного байесовского вывода. В [12] приводится алгоритм, основанный на использовании случайных марковских полей для моделирования изображений и метода постепенной невыпуклой оптимизации (GNC) для получения изображения ВР.

Наконец, в статьях [13, 14] авторами данной работы рассматривался алгоритм, основанный на итеративной обработке поступающих кадров НР с реализацией схемы оптимальной квазилинейной фильтрации и суперпиксельного представления изображений. В [15] задача итеративного восстановления изображения ВР решалась авторами при помощи нейронных сетей. Полученные в этих работах результаты позволяют утверждать, что потенциальные возможности итеративной обработки последовательности кадров на основе схемы оптимальной фильтрации с выполнением суперпиксельной сегментации АП на каждом кадре достаточно велики в плане достижения высококачественных получаемых изображений.

Тем не менее остается открытым вопрос об адекватности используемых в рамках этих работ математических моделей и, в частности, условно-гауссовской модели данных для описания обрабатываемых изображений. Поэтому целью данной работы является исследование возможностей построения алгоритмов многокадрового сверхразрешения в условиях АП, реализующих общую схему итеративной обработки изображений НР с сегментацией АП на входных изображениях, полностью основанную на использовании глубоких сверточных нейронных сетей.

2. Постановка задачи и общая схема обработки изображений

Пусть имеется T изображений НР I_t размерности (h^L, w^L, c) , которые могут быть подвержены воздействию АП. При этом в общем виде свойства АП и расположение ЛОЗ не известны. Требуется получить одно изображение I^H разрешения выше исходного (h^H, w^H, c) , $h^L < h^H$, $w^L < w^H$ и не подверженное воздействию АП.

Для решения этой задачи предлагается использовать итеративный подход к построению многокадрового СР, который ранее был реализован в работах авторов [13, 14] при построении рекуррентных алгоритмов оптимальной линейной и нелинейной фильтрации.

Общая схема обработки при реализации такого подхода дана на рис. 1 и представляется разумной и для построения алгоритмов машинного обучения. Она состоит из двух основных шагов.

1. Регистрация очередного изображения (кадра) НР и оценка смещения между двумя соседними в последовательности кадрами.
2. Обновление ранее полученной оценки изображения ВР путем использования новой информации.

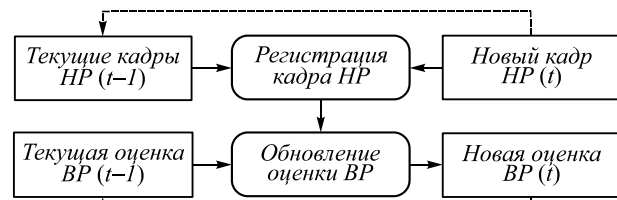


Рис. 1. Общая схема итеративного уточнения оценки изображения ВР

Достоинством данной схемы является то, что процедура оценки смещения камеры между изображениями позволяет повысить точность итоговой оценки изображения ВР, а накопление информации, полученной по нескольким кадрам НР, помогает устранить негативное воздействие АП одновременно с процедурой повышения разрешения. В пользу такой схемы обработки говорит также то, что сама последовательность кадров НР может иметь произвольную длину, при этом каждый кадр обрабатывается один раз.

В [13, 14] было показано, что использование данной схемы обработки для реализации алгоритма оптимальной линейной и квазилинейной фильтрации изображений НР превосходит схожие алгоритмы [10–12] по точности восстановления изображения ВР. Тем не менее данный подход имеет и ряд недостатков:

- использование математических моделей, опирающихся на не всегда адекватные реалиям предположения и приближения относительно статистических свойствах исходных изображений и возникающих ложных наблюдений;
- достаточно длительное время, затрачиваемое на обработку последовательности изображений.

3. Предлагаемый подход для обработки изображения на основе глубоких нейронных сетей

Использование искусственных нейронных сетей для решения задач обработки изображений в ходе итеративного построения многокадрового СР позволяет решить обе проблемы рассмотренного выше алгоритма [13], так как их применение не предполагает использования моделей данных, а быстроедействие уже обученной нейронной сети существенно выше быстрогодействия оптимального линейного фильтра.

В [15] авторами был реализован итеративный подход к многокадровому СР, использующий сверточные нейронные сети как для регистрации изображений, так и для повышения их разрешения, основываясь на алгоритме FVRSR [3]. Для регистрации изображений НР использовалась сеть LiteFlowNet2 [16], а для повышения их детализации использовалась

модифицированная сеть SRNet из [3]. Нейронная сеть для повышения качества изображения принимает на вход два изображения: текущую оценку ВР и очередное изображение НР, разрешение которого было повышено при помощи бикубической интерполяции. На выходе получается одно изображение, являющееся обновлённой оценкой ВР.

Результаты проведённого эксперимента показали, что такой подход частично позволяет справиться с устранением АП на изображениях НР. Однако в некоторых случаях области получаемого изображения ВР, изначально подверженные воздействию АП, могут быть несколько искажены графическими артефактами. Потенциальным улучшением данной схемы будет являться добавление в неё этапов расчёта усреднённой начальной оценки ВР, а главное, независимой сегментации ложных наблюдений на каждом изображении НР, которые обрабатываются в последовательном режиме.

Рассмотрим реализацию итеративной обработки по схеме рис. 1, комбинирующую сильные стороны перечисленных алгоритмов, такие как использование усреднённой начальной оценки изображения ВР и процедуры сегментации областей, подверженных АП. Такая схема обработки изображений по своей сути эквивалентна [13], но все основные этапы обработки реализуются при помощи нейронных сетей по аналогии с [15]. При этом в ней на отдельных этапах используются как предварительно обученные нейронные сети, так и особая архитектура DAG-сети, отвечающая за формирование новой оценки ВР и обученная в ходе наших экспериментов.

Предлагаемая схема обработки описывается следующим алгоритмом (рис. 2).

1. На основе входных кадров НР происходит формирование исходной оценки изображения ВР (блок 1 на рис. 2).
2. Осуществляется регистрация очередного кадра НР (блок 2).
3. Происходит сегментация областей, подверженных воздействию АП, на очередном кадре НР (блок 3).
4. На основе текущей оценки изображения ВР, очередного зарегистрированного кадра НР и результатов его сегментации осуществляется обновление оценки ВР (блок 4).
5. Шаги 2–4 повторяются, пока не будет обработано требуемое количество кадров НР.
6. Последняя полученная оценка ВР становится выходным изображением ВР.

Рассмотрим шаги данного алгоритма более подробно.

Вычисление начальной оценки ВР (блок 1) осуществляется аналогично [13] путём повышения разрешения каждого кадра НР при помощи бикубической интерполяции и усреднения полученных изображений ВР. Данная процедура позволяет существенно сократить негативное воздействие АП.

ражений ВР. Данная процедура позволяет существенно сократить негативное воздействие АП.

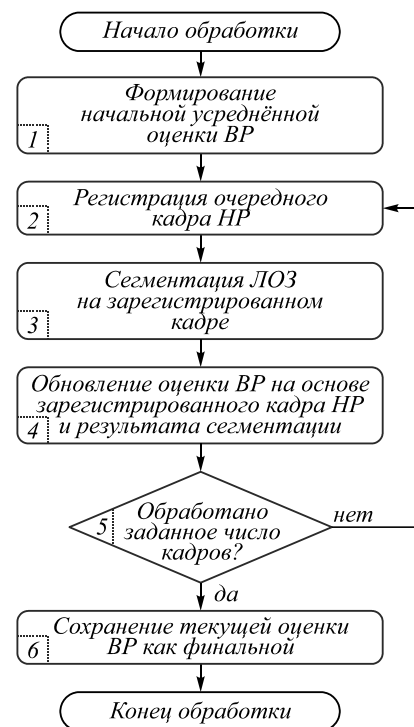


Рис. 2. Предлагаемый алгоритм обработки последовательности изображений НР с целью повышения разрешения

Регистрация очередного изображения НР (блок 2) происходит аналогично [15]: оптический поток между первым и текущим кадрами НР оценивается при помощи нейронной сети LiteFlowNet2 и затем используется для наложения текущего кадра на первый. LiteFlowNet2 реализует пирамидальный подход к получению оптического потока с переменным числом используемых свёрточных слоёв. При этом каждый уровень последовательно повышает детализацию результата и состоит из двух подсетей, работающих и обучаемых совместно.

Сегментация ЛОЗ на изображении НР (блок 3) осуществляется при помощи нейронной сети FCN-AlexNet [15], которая осуществляет семантическую сегментацию пикселей входного изображения на два класса: полезные и ложные наблюдения. FCN-AlexNet получается путём трансформации нейронной сети для классификации изображений AlexNet [18], в ходе которой осуществляется замена полносвязных слоёв, а свёрточные слои остаются без изменений (рис. 3). Так же, как и в [13], полученная в ходе сегментации информация используется в дальнейшем с целью исключения ложных наблюдений (АП) при обновлении оценки изображения ВР на следующем шаге.

Формирование новой оценки ВР (блок 4, рис. 2) осуществляется при помощи предложенной глубокой DAG-сети, состоящей из двух параллельно включённых свёрточных сетей, объединяемых в выходной ча-

сти для получения обновлённой оценки ВР. Архитектура этой сети представлена на рис. 4. Использование свёрточных слоёв является стандартным подходом для задачи СР.

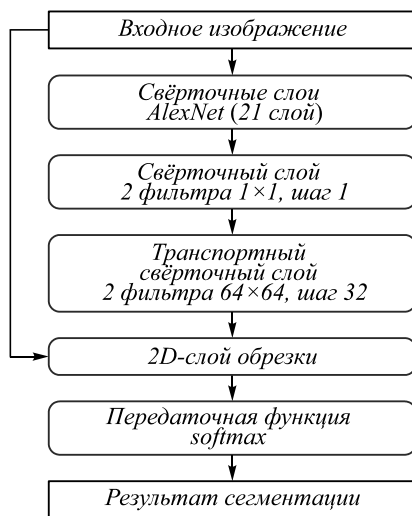


Рис. 3. Архитектура нейронной сети FCN-AlexNet для семантической сегментации на два класса

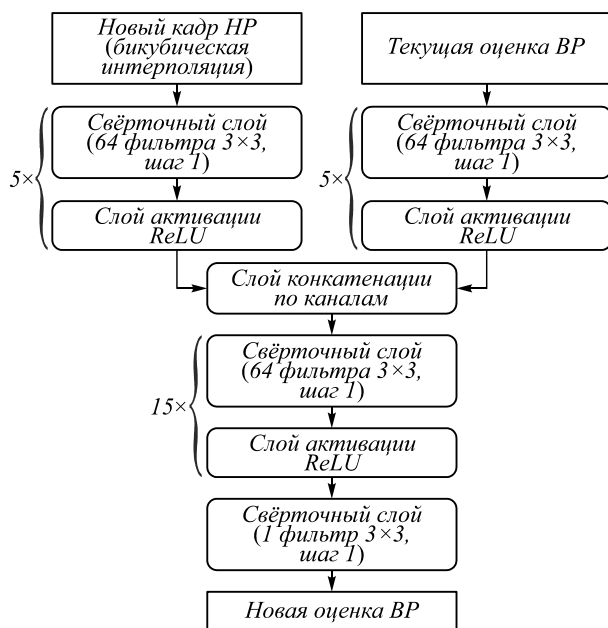


Рис. 4. Архитектура нейронной DAG-сети для обновления оценки изображения ВР

На вход DAG-сети подаются следующие изображения:

- сформированная к текущему шагу оценка изображения ВР;
- очередной кадр НР с удалёнными ЛОЗ, разрешение которого было повышено до целевого.

Как в [19], на вход данной нейронной сети подаются цветовые компоненты яркости Y изображений, представленных в пространстве $YCbCr$. На выходе нейронной сети получается одноканальное изображение ВР, которое необходимо сложить с компонентой

яркости текущей оценки ВР, чтобы получить обновлённую компоненту яркости оценки ВР.

Таким образом, процесс обновления оценки ВР с учётом результатов сегментации ЛОЗ состоит из следующих этапов, представленных на схеме на рис. 5.

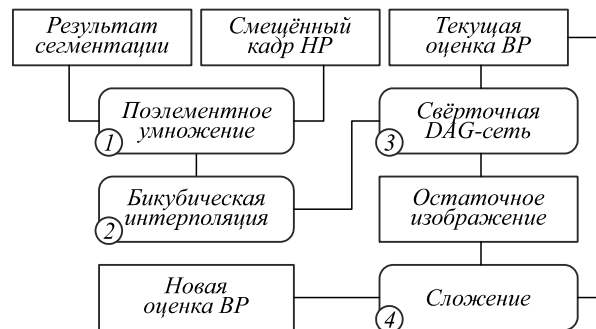


Рис. 5. Схема этапа расчёта новой оценки изображения ВР

1. Удаление областей, подверженных воздействию АП, путём поэлементного умножения пикселей очередного изображения ВР на бинарную маску, полученную в ходе его сегментации (блок 1 на рис. 5).
2. Повышение полученного разрешения до целевого при помощи бикубической интерполяции (блок 2).
3. Получение остаточного изображения при помощи DAG-сети (блок 3).
4. Сложение остаточного изображения с компонентой яркости текущей оценки ВР (блок 4).

На рис. 6а–е как пример приведены изображения, иллюстрирующие различные этапы обработки кадра НР.

Таким образом, в рамках общей схемы итеративной обработки предложен комплексированный алгоритм построения многокадрового сверхразрешения, основанный на использовании глубоких нейронных сетей на всех этапах обработки, включая регистрацию изображений, сегментацию областей, подверженных влиянию АП, и преобразования, непосредственно направленные на повышение качества изображений.

4. Результаты эксперимента и их обсуждение

Для обучения рассмотренной выше нейронной DAG-сети для итеративного обновления оценки ВР использовалась обучающая выборка из [15], включающая в себя 40000 наборов цветных кадров размером 64×64 , состоящих из 10 изображений каждый (рис. 7а). Исходные изображения (рис. 7б) были получены из набора SUN [20], а затем каждый набор кадров НР был создан путём случайных смещений и поворотов исходных изображений, к которым затем были случайным образом применены АП (при этом некоторые кадры или целые последовательности кадров могли остаться неискаженными). Для моделирования АП был использован предложенный в [21] алгоритм, формирующий поток ЛОЗ со случайным начальным положением и формой. Разрешение изоб-

ражений в обучающей выборке было понижено в 2 раза путём прореживания исходных изображений.

Обучение DAG-сети проводилось при помощи алгоритма Adam.

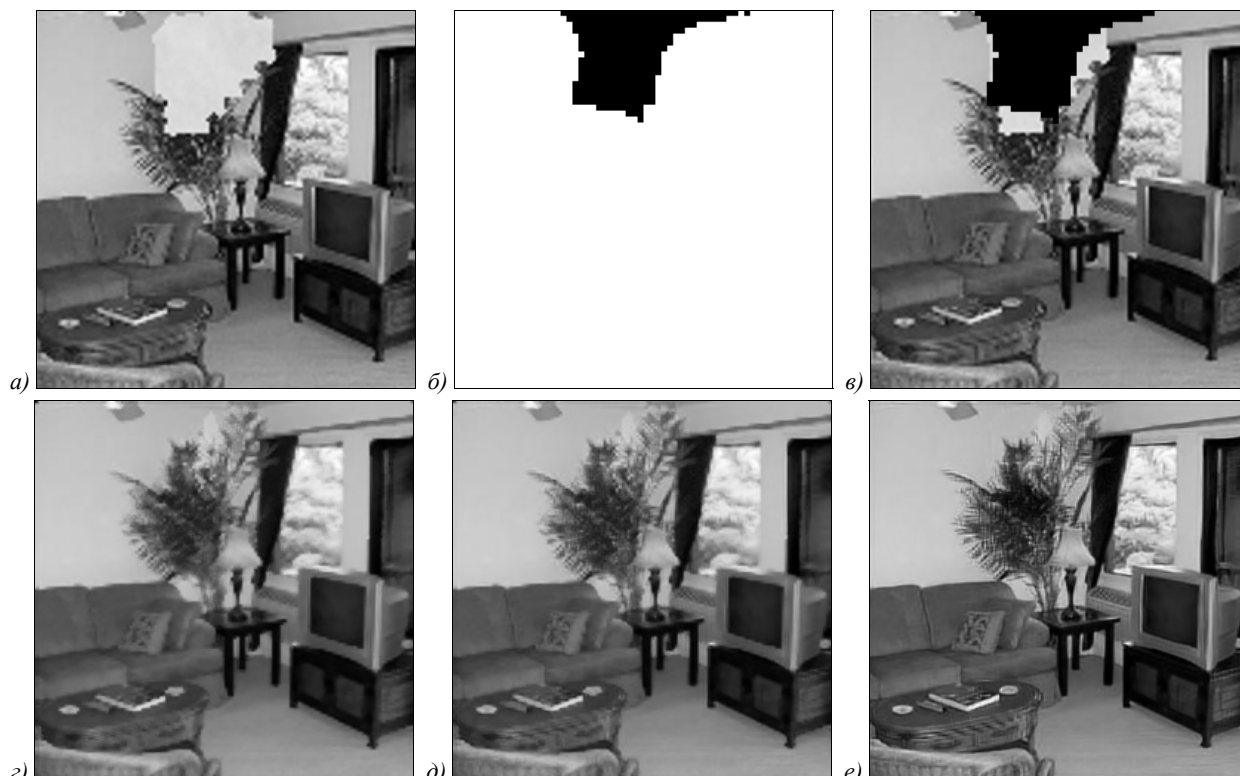


Рис. 6. Изображения на различных этапах обработки: первый смещённый кадр HP (бикубическая интерполяция) (а); результат семантической сегментации первого кадра (бикубическая интерполяция) (б); первый смещённый кадр с удалёнными ЛОЗ (бикубическая интерполяция) (в); начальная оценка BP (г); оценка BP после обработки первого HP (д); оценка BP после обработки всех кадров HP (е)

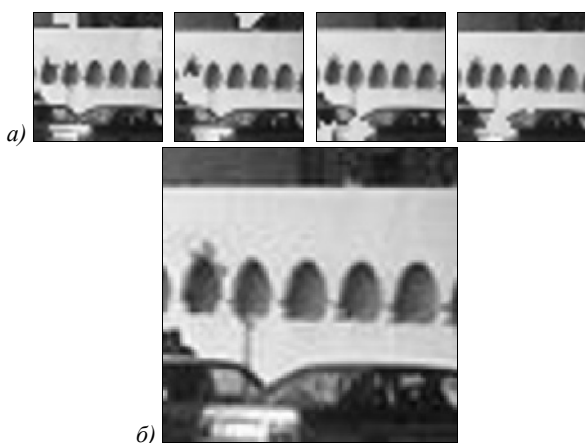


Рис. 7. Примеры изображений HP из входной последовательности кадров (а) и соответствующего им изображения BP(б) из обучающей выборки для нейронной DAG-сети, формирующей обновление оценки BP

Для дообучения нейронной сети для семантической сегментации FCN-AlexNet, полученной на основе предварительно обученной AlexNet, использовались сведения о ЛОЗ, полученные при формировании обучающих данных для DAG-сети (рис. 8). Таким образом, обучающая выборка состояла из 400 тысяч цветных изображений и соответствующих им бинарных масок ЛОЗ. Обучение проводилось при

помощи алгоритма стохастического градиентного спуска с импульсом (SGDM).

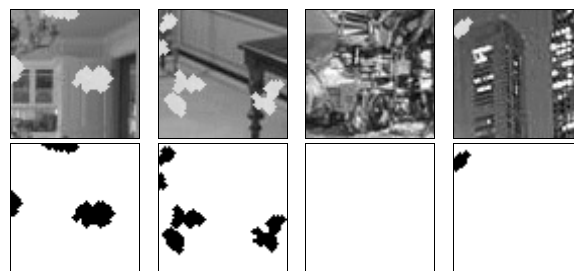


Рис. 8. Примеры изображений HP и соответствующие им ЛОЗ из обучающей выборки для FCN-AlexNet

Для оценки оптического потока между кадрами использовалась предварительно обученная её авто-рами сеть LiteFlowNet2.

Для тестирования предлагаемого алгоритма и его сравнения с другими алгоритмами использовались 100 наборов из 10 кадров размером 128×128 , полученных таким же образом, как и обучающая выборка. При этом рассматривались изображения как подверженные воздействию АП, так и не искажённые ложными наблюдениями.

В рамках экспериментального исследования было проведено сравнение предложенного подхода со следующими алгоритмами многокадрового СР в

условиях наличия АП на входных изображениях: алгоритм, основанный на методе спиновых стёкол [10], алгоритм, основанный на применении модели случайного марковского поля [12], алгоритм, основанный на оптимальной линейной фильтрации [13], алгоритм, реализованный авторами и основанный на использовании нейронных сетей без выполнения этапа сегментации АП [15]. Кроме того, в рамках экспериментального исследования был рассмотрен алгоритм многокадрового СР FRVSR [3], также основанный на использовании свёрточных нейронных сетей, но без учёта специфики АП. Как и предлагаемый авторами алгоритм, FRVSR реализует итеративный подход к получению изображения ВР (рис. 1) и является его наиболее близким аналогом.

На рис. 9 приведён пример оригинального изображения ВР (256 × 256), наряду с восстановленными изображениями ВР, полученными при помощи различных алгоритмов. Визуальный анализ входных и выходных изображений показывает, что предложенный алгоритм выполняет устранение АП и одновременное восстановление детализации изображений на одном уровне с результатами, полученными в [13], и с существенно лучшим качеством, чем результаты, полученные в [3, 10, 12, 15].

Для численного сравнения полученных в ходе экспериментов результатов использовались традиционные показатели оценки качества восстанавливаемых изображений:

- пиковое отношение сигнала к шуму (PSNR), задающееся на основе отношения максимального значения пикселей в изображении к среднеквадратической ошибке (MSE) исходного и восстановленного изображений;
- индекс структурного сходства (SSIM), учитывающий, в отличие от PSNR и MSE, «восприятие ошибки» благодаря анализу структурных изменений свойств исходного и восстановленного изображений [22].

Чем больше значения данных критериев, тем меньше разница между исходными и полученными на выходе изображениями ВР и лучше результат построения СР. Усреднённые значения этих показате-

лей, полученные в ходе эксперимента, приведены в табл. 1, наряду со средним временем обработки одной последовательности изображений. Её анализ подтверждает результаты качественного визуального сравнения формируемых изображений ВР: предложенная схема построения СР восстанавливает изображение ВР с большей точностью по обоим введённым показателям, чем оптимальный квазилинейный фильтр [13], имея, наряду с [15], существенно лучшее быстродействие. Последнее преимущество достигается за счет того, что в отличие от алгоритма [13] рассматриваемый подход не требует сложных матричных вычислений. Кроме того, рассмотренный алгоритм имеет точность одного порядка с алгоритмом [3] в случае отсутствия АП на входных изображениях. При этом он значительно превосходит алгоритм [3] в случае наличия АП. Следует отметить, что все рассмотренные алгоритмы справляются с задачей восстановления изображения ВР в случае отсутствия АП на входных изображениях.

Также стоит отметить тот факт, что замена отдельных используемых алгоритмов может привести к большей точности обработки. В частности, в работе используется предварительно обученная нейронная сеть для определения оптического потока между двумя изображениями, которая не учитывает специфику, связанную с появлением АП, что потенциально может повысить результирующее качество обработки.

Заключение

Данная статья посвящена алгоритмам построения многокадрового СР в условиях аномальных наблюдений. Проведён анализ предшествующих работ, включая и работы авторов, опирающиеся на использование методов оптимальной фильтрации к построению многокадрового сверхразрешения.

В рамках указанной схемы использован итеративный подход к решению задачи построения СР при обработке последовательности кадров НР с одновременным устранением воздействия действующих на них АП, базирующийся на использовании методов и технологий машинного обучения.

Табл. 1. Численные результаты эксперимента

	Алгоритм на основе модели спиновых стёкол [10]	Алгоритм на основе модели случайного марковского поля [12]	Оптимальный условно-линейный фильтр [13]	Нейронные сети (FRVSR) [3]	Нейронные сети без сегментации ЛОЗ [15]	Предлагаемый комплексированный алгоритм
PSNR (дБ) с АП	21,06	21,28	25,16	24,14	21,63	29,11
SSIM с АП	0,89	0,89	0,93	0,88	0,85	0,97
PSNR (дБ) без АП	21,22	21,53	25,17	30,70	22,18	30,38
SSIM без АП	0,89	0,90	0,93	0,98	0,87	0,98
Время (с)	545	452	223	10	10	9

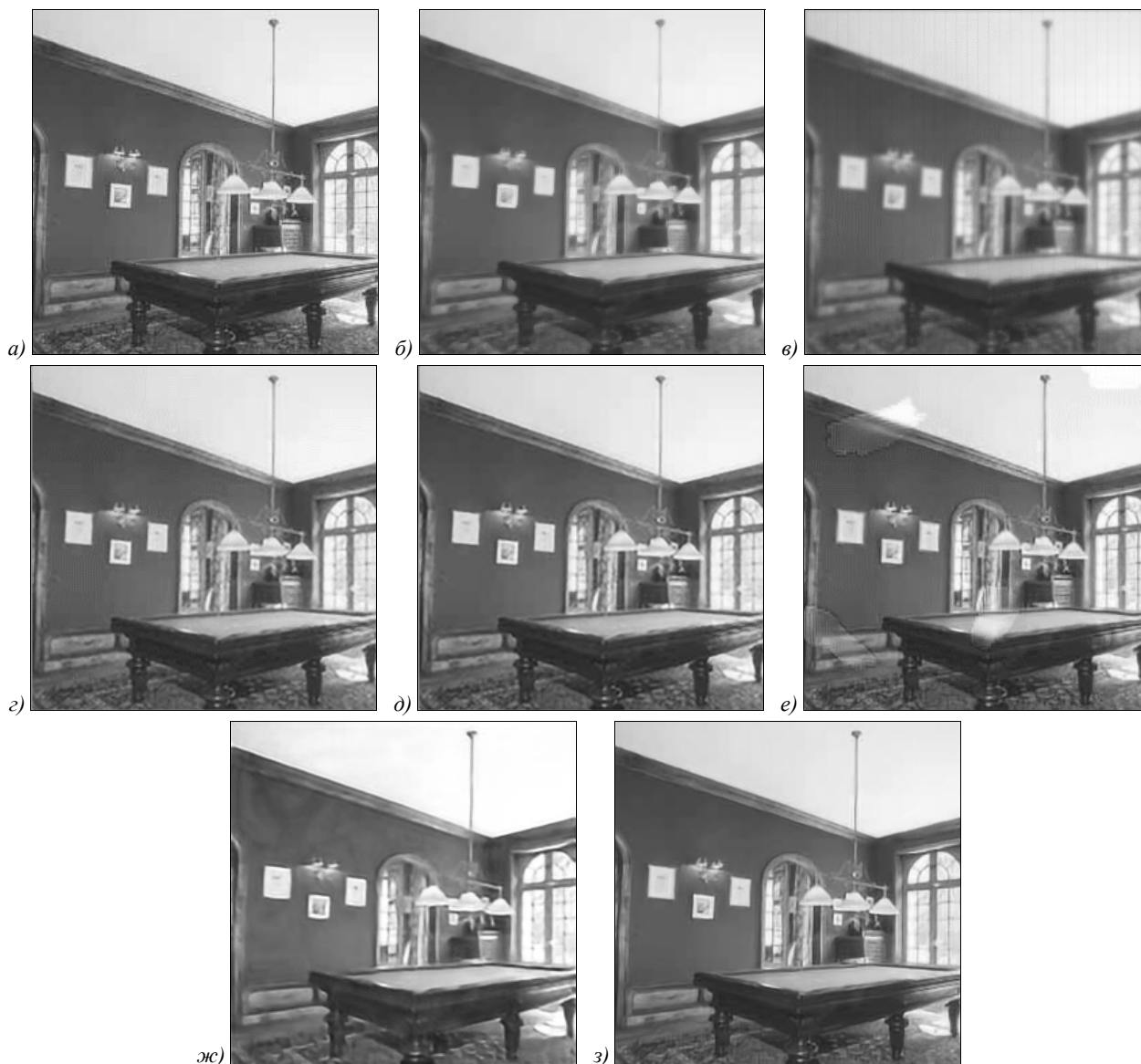


Рис. 9. Примеры изображений ВР: оригинальное изображение ВР (а); усреднённая начальная оценка (б); результат алгоритма, использующего модель спиновых стёкол [10] (в); результат алгоритма, использующего модель случайных марковских полей [12] (г); результат алгоритма, основанного на оптимальной условно-линейной фильтрации [13] (д); результат алгоритма, основанного на нейронных сетях без учёта АП [3] (е); результат алгоритма, основанного на нейронных сетях без сегментации ЛОЗ [15] (ж); результат рассматриваемого алгоритма, основанного на нейронных сетях с сегментацией ЛОЗ (з)

Предложен и исследован новый комплексный алгоритм построения многокадрового сверхразрешения, основанный на использовании глубоких свёрточных нейронных сетей на всех этапах обработки, включая регистрацию изображений, сегментацию областей, подверженных влиянию АП, и обработку изображений, направленную непосредственно на повышение их качества. В рамках этого алгоритма предложена искусственная нейронная сеть с архитектурой в виде направленного ациклического графа, используемой для обновления оценки изображения ВР.

В ходе исследования установлено, что применение предлагаемого комплексного алгоритма обладает определёнными преимуществами по сравнению с рассмотренными известными аналогами как в плане до-

стигаемой точности обработки, так и в плане повышения её быстродействия. Так, являясь на порядок более быстрым по сравнению с алгоритмом, основанным на оптимальной линейной фильтрации, предлагаемый алгоритм имеет большую точность (в среднем на 4 дБ по метрике PSNR и на 0,04 по метрике SSIM). По сравнению с известным алгоритмом, основанным на использовании свёрточных нейронных сетей и синтезированным без учёта специфики обработки изображений в условиях АП, рассматриваемый алгоритм имеет схожую точность при обработке изображений, не подверженных воздействию АП. При этом предложенный алгоритм превосходит известный по всем показателям качества обработки в случае наличия ложных наблюдений в виде АП на входных изображениях.

В качестве направлений дальнейших исследований целесообразно рассмотреть возможность замены отдельных предобученных моделей на более точные аналоги, лучше учитывающие специфику воздействия АП, организацию совместного обучения используемых нейронных сетей, а также вопросы использования предложенных алгоритмов для обработки видеопоследовательностей.

References

- [1] Villegas OV, Nandayapa M, Solo I, eds. Advanced topics on computer vision. Springer; 2018.
- [2] Huang D, Liu H. A short survey of image super resolution algorithms. Journal of Computer Science Technology Updates 2015; 2(2): 19-29. DOI: 10.15379/2410-2938.2015.02.02.03.
- [3] Sajjadi M, Vemulapalli R, Brown M. Frame-recurrent video super-resolution. IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2018: 6626-6634. DOI: 10.1109/CVPR.2018.00693.
- [4] Kawulok M, Benecki P, Piechaczek S, Hrynczenko K, Kostrzewa D, Nalepa J. Deep learning for multiple-image super-resolution. IEEE Geosci Remote Sens Lett 2019; 17(6): 1062-1066. DOI: 10.1109/LGRS.2019.2940483.
- [5] Li Z, Li S, Wang J, Wang H. A novel multi-frame color images super-resolution framework based on deep convolutional neural network. Proc 5th Int Conf on Measurement, Instrumentation and Automation (ICMIA) 2016: 634-639. DOI: 10.2991/icmia-16.2016.115.
- [6] Kappeler A, Yoo S, Dai Q, Katsaggelos AK. Video super-resolution with convolutional neural networks. IEEE Trans Comput Imaging 2016; 2(2): 109-122. DOI: 10.1109/TCI.2016.2532323.
- [7] Molini AB, Valsesia D, Fracastoro G, Magli E. DeepSUM: Deep neural network for super-resolution of unregistered multitemporal images. IEEE Trans Geosci Remote Sens 2020; 58(5): 3644-3656. DOI: 10.1109/TGRS.2019.2959248.
- [8] Kim J, Lee JK, Lee KM. Deeply-recursive convolutional network for image super-resolution. IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2016: 1637-1645. DOI: 10.1109/CVPR.2016.181.
- [9] Deudon M, Kalaitzis A, Goytom I, Arefin R, Lin Z, Sankaran K, Michalski V, Kahou S, Cornebise J, Bengio Y. HighRes-net: Recursive fusion for multi-frame super-resolution of satellite imagery. arXiv Preprint arXiv:2002.06460 2020. Source: (https://arxiv.org/abs/2002.06460).
- [10] Kanemura A, Maeda S, Ishii S. Image superresolution under spatially structured noise. 2007 IEEE Int Symposium on Signal Processing and Information Technology 2007: 275-280. DOI: 10.1109/ISSPIT.2007.4458156.
- [11] Fukuda W, Kanemura A, Maeda S, Ishii S. Superresolution from occluded scenes. In Book: Leung CS, Lee M, Chan JH, eds. ICONIP '09: Proceedings of the 16th International Conference on Neural Information Processing 2009, II: 19-27. DOI: 10.1007/978-3-642-10684-2_3.
- [12] Negi CS, Mandal K, Sahay RR, Kankanhalli MS. Super-resolution de-fencing: Simultaneous fence removal and high-resolution image recovery using videos. Proc IEEE Int Conf on Multimedia and Expo Workshops 2014: 1-6. DOI: 10.1109/ICMEW.2014.6890641.
- [13] Savvin SV, Sirorta AA, Ivankov AYu. Methods and algorithms of super-resolution for image sequence under applicative noise [In Russian]. Izvestiya RAN. Control Theory and Systems 2021; 3: 136-148. DOI: 10.31857/S0002338821030070.
- [14] Savvin SV, Sirota AA. Performance of super-resolution algorithms under applicative noise. J Phys Conf Ser 2019; 1479: 11-13. DOI: 10.1088/1742-6596/1479/1/012080.
- [15] Savvin SV, Sirota AA. An algorithm for multi-frame image super-resolution under applicative noise based on a convolutional neural network. 2nd Int Conf on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA) 2020: 422-424. DOI: 10.1109/SUMMA50634.2020.9280698.
- [16] Hui TW, Tang X, Loy CC. A lightweight optical flow cnn – revisiting data fidelity and regularization. IEEE Trans Pattern Anal Mach Intell 2020; 43(8): 2555-2569. DOI: 10.1109/TPAMI.2020.2976928.
- [17] Balas VE, Roy SS, Sharma D, Samui P, eds. Handbook of deep learning applications. Springer; 2019.
- [18] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM 2017; 60(6): 84-90. DOI: 10.1145/3065386.
- [19] Kim J, Lee JK, Lee KM. Accurate image super-resolution using very deep convolutional networks. IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2016: 1646-1654. DOI: 10.1109/CVPR.2016.182.
- [20] Xiao J, Hays J, Ehinger K, Oliva A, Torral A. SUN database: Large-scale scene recognition from Abbey to Zoo. IEEE Conf on Computer Vision and Pattern Recognition 2010: 3485-3492. DOI: 10.1109/CVPR.2010.5539970.
- [21] Alginov EK, Sirota AA. Analysis and computer modeling of informational processes and systems [In Russian]. Moscow: "Dialog-MIFI: Publisher; 2009.
- [22] Wang Z, Bovik AC, Sheikh AR, 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.

Сведения об авторах

Саввин Сергей Викторович, 1993 года рождения, в 2017 году окончил Воронежский государственный университет по направлению 09.04.02 «Информационные системы и технологии». Аспирант кафедры технологий обработки и защиты информации Воронежского государственного университета. Область научных интересов: обработка цифровых изображений, программирование. E-mail: savvin_s_v@sc.vsu.ru.

Сирота Александр Анатольевич, 1954 года рождения, в 1976 году окончил Воронежский государственный университет по специальности «Радиофизика и электроника». Доктор технических наук (1995 год), профессор, заведует кафедрой технологий обработки и защиты информации Воронежского государственного университета. Область научных интересов: синтез и анализ систем сбора и обработки информации, методы и технологии компьютерного моделирования информационных процессов и систем, системный анализ в сфере информационной безопасности, компьютерная обработка изображений, нейронные сети и нейросетевые технологии в системах принятия решений. E-mail: sir@cs.vsu.ru.

ГРНТИ: 28.23.15

Поступила в редакцию 7 апреля 2021 г. Окончательный вариант – 1 июля 2021 г.

Algorithms for multi-frame image super-resolution under applicative noise based on deep neural networks

S.V. Savvin¹, A.A. Sirota¹

¹ Voronezh State University, 394018, Voronezh, Russia, Universitetskaya pl.1

Abstract

The article describes algorithms for multi-frame image super-resolution, which recover high-resolution images from a sequence of low-resolution images of the same scene under applicative noise. Applicative noise generates local regions of outlying observations in each image and reduces the image resolution. So far, little attention has been paid to this problem. At the same time, the use of deep neural networks is considered to be a promising method of image processing, including multi-frame image super-resolution. The article considers the existing solutions to the problem and suggests a new approach based on using several pre-trained convolutional neural networks and directed acyclic graph neural networks trained by the authors. The developed approach and the algorithms based on this approach involve iterative processing of the input sequence of low-resolution images using different neural networks at different processing stages. The stages include registration of low-resolution images, their segmentation performed in order to determine regions damaged by applicative noise, and transformation performed in order to increase the resolution. The approach combines the strengths of the existing solutions while lacking their drawbacks resulting from the use of approximate mathematical data models required for the synthesis of the image processing algorithms within the statistical theory of solutions. The experimental studies demonstrated that the suggested algorithm is fully functional and allows more accurate recovery of high-resolution images than the existing analogues.

Keywords: digital image processing, multi-frame superresolution, convolutional neural networks, deep learning, applicative noise.

Citation: Savvin SV, Sirota AA. Algorithms for multi-frame image super-resolution under applicative noise based on deep neural networks. *Computer Optics* 2022; 46(1): 130-138. DOI: 10.18287/2412-6179-CO-904.

Authors' information

Sergey Viktorovich Savvin (b. 1993) graduated from Voronezh State University in 2017, majoring in Information Systems and Technologies. Currently a postgraduate student of Information Processing and Security Technologies department at Voronezh State University. Research interests: processing of digital images, programming. E-mail: savvin_s_v@sc.vsu.ru.

Alexander Anatolievich Sirota (b. 1954) graduated from Voronezh State University in 1976 majoring in Radiophysics and Electronics. Professor, Doctor of Technical Sciences (since 1995). Currently head of Information Processing and Security Technologies department at Voronezh State University. Research interests: analysis and design of information collection and processing systems, methods and techniques of information processes and systems computer modeling, system analysis in information security, digital image processing, neural networks and neural network technologies in decision-making systems. E-mail: sir@cs.vsu.ru.

Received April 7, 2021. The final version – July 1, 2021.
