

УДК 519.688

А.А. ЛУКЪЯНИЦА, доктор
технических наук
В.Б. ЭПЕНЕТОССКИЙ, кандидат
технических наук, старший
научный сотрудник
В.Н. ИСАЕВ

ВЫСОКОЭФФЕКТИВНЫЙ АЛГОРИТМ ПОВЫШЕНИЯ ДЕТАЛЬНОСТИ ОПТИКО-ЭЛЕКТРОННЫХ ИЗОБРАЖЕНИЙ, ПОЛУЧАЕМЫХ ЛЕТАТЕЛЬНЫМИ АППАРАТАМИ

В настоящей работе представлен алгоритм повышения детальности (разрешения на местности) оптико-электронных изображений, получаемых с использованием летательных аппаратов воздушного и космического базирования, отличающийся от существующих тем, что он обладает высоким быстродействием, требует минимальных вычислительных ресурсов и может быть реализован аппаратно на существующей отечественной элементной базе. Это позволяет реализовать цифровую обработку оптико-электронных изображений непосредственно на борту летательного аппарата в реальном масштабе времени. Проведенное сравнение представленного алгоритма с существующими показало более чем двукратное повышение точности при получении изображений высокого разрешения.

Ключевые слова: летательные аппараты; дистанционное зондирование земли; суперразрешение; связательная нейронная сеть; метод регуляризации; распараллеливание.

Введение

Основной тенденцией развития систем дистанционного зондирования Земли (ДЗЗ) воздушного и космического базирования, во все периоды их существования и развития, является постоянное улучшение разрешающей способности оптико-электронной аппаратуры наблюдения. В техническом плане повышение разрешающей способности реализуется через увеличение диаметра и фокусного расстояния используемой оптической системы; уменьшение размера пикселей фотоприемных матриц, используемых в аппаратуре; снижение высоты рабочей орбиты (касается космических аппаратов (КА) ДЗЗ). Зачастую на практике все три способа используются одновременно. Одновременно с разрешающей способностью одной из важнейших характеристик

оптических систем высокого разрешения является полоса захвата. Этот параметр напрямую связан с производительностью процесса съемки (снимаемой площадью и количеством получаемой информации). При этом ширина захвата определяется рабочим полем зрения оптической системы камеры, а также общей длиной строки фотоприемника. Для объективной оценки информативности оптической системы принципиальное значение имеет даже не сама полоса захвата, а ее отношение к пространственному разрешению (проекции пикселя на Землю) [1].

Вместе с улучшением разрешающей способности, ростом длины строки и информационной производительности, также увеличивается и масса оптико-электронной аппаратуры и летательного аппарата в целом, а, следовательно, и стоимость всего комплекса ДЗЗ. В этих условиях (в части КА ДЗЗ) наблюдаются два пути развития. Первый – это создание дорогих и тяжелых КА, на которых одновременно реализуются максимальные характеристики по основным потребительским параметрам. Данное направление требует больших объемов финансирования, длительного времени разработки, наличия отечественной элементной базы и научно-технического задела. Второй – создание относительно недорогих малых космических аппаратов (МКА), в которых разработчики не стремятся реализовать предельные характеристики своих систем сразу по всем направлениям (разрешение, производительность, точность), а сосредотачиваются только на одном или нескольких из этих параметров, реализуя необходимую для себя и коммерчески привлекательную для потребителей систему наблюдения. Однако МКА ДЗЗ на сегодняшний день имеет технологические ограничения разрешающей способности примерно на уровне 1 м [2]. Приближение к этому пределу отечественных проектов МКА может повлечь за собой серьезный рост затрат на их реализацию из-за отсутствия научно-технологического задела по нетрадиционным материалам, отсутствия отечественных аналогов современной высокоинтегрированной элементной базы или невозможностью ее закупки за рубежом.

В этих условиях представляется целесообразным третий путь – использование относительно недорогих, отработанных КА с оптико-электронной аппаратурой (например, «Канопус-В») с разрешением порядка 2 метров, а разрешающую способность на местности повышать за счет последующей компьютерной обработки оптико-электронных изображений. Повышение разрешения путем компьютерной обработки в мировой

научной литературе называют сверхразрешением или суперразрешением (англ. Super Resolution). Если ввести английский термин в качестве запроса в поисковой системе Google, то получим более 3,1 млн. ссылок на статьи по этой тематике, что говорит о большом интересе исследователей к данной проблеме. Здесь следует выделить два направления исследований: увеличение разрешения единственного снимка либо использование нескольких снимков одного и того же изображения при небольшом сдвиге камеры (которое часто называют съемкой с перекрытием). Львиная доля разработок посвящена обработке одного снимка ввиду большого числа областей приложения. При решении этой задачи между имеющимися пикселями вставляются дополнительные, интенсивность которых определяют методом интерполяции. В простейшем случае используют метод ближайших соседей [3], билинейную [4] либо бикубическую интерполяцию [5]. Последний метод, в частности, используется в известной программе Adobe Photoshop. В этих методах восполнение всех промежуточных пикселей проводится по фиксированному алгоритму. Несмотря на то, что разработка и применение перечисленных алгоритмов началась во второй половине прошлого века, интерес к их совершенствованию постоянно растет, что, в частности, видно из приведенных ссылок.

Более продвинутыми являются методы, основанные на применении искусственного интеллекта, в основе которых лежит идея восполнения недостающих пикселей нейронной сетью, которая была обучена на большом числе изображений. Например, в широко цитируемых работах [6; 7] использовалась предварительно обученная сверточная сеть, в [8] – рекурсивная сверточная сеть, в [9; 10] – генеративно-состязательные нейросети. Применение этих методов позволяет получить лучшие результаты, чем использование интерполяции, поскольку нейросеть обладает способностью адаптироваться к конкретному фрагменту изображения и, как следствие, проводить более качественное восполнение недостающих пикселей.

В последнее время большую популярность обрел метод RAISR [11], разработка которого началась сотрудниками компании Google и в настоящее время часто используется для сокращения объема при хранении и передаче изображений. Этот метод сочетает в себе большинство позитивных качеств вышеупомянутых методов: он анализирует окрестность каждого восстанавливаемого пикселя и подбирает для него один из предварительно обученных фильтров, устраняя при этом различные артефакты типа

алиасинга. Поскольку обучение фильтров, хоть и занимает значительное время, проводится на изображениях обрабатываемого класса заранее, и с учетом разумно организованного хэширования позволяет проводить обработку на один-два порядка быстрее других известных методов [11].

Следует отметить, что несмотря на все кратко упомянутое разнообразие, по своей сути способы использования единственного изображения для повышения разрешения позволяют лишь назначить восстанавливаемым пикселям наиболее статистически достоверное значение, которое с наибольшей вероятностью появлялось в аналогичных фрагментах изображения, либо является интерполяцией соседних значений. В то же время методы, основанные на повышении разрешения из последовательности изображений с недостаточной дискретизацией, могут использовать преимущества дополнительных пространственно-временных данных, имеющихся в этой последовательности. В частности, движение камеры относительно сцены позволяет получить кадры, содержащие схожую, но не идентичную информацию. Благодаря этому возникает возможность использовать эту информацию для реконструкции более мелких деталей изображения, используя принципы, на которых основана, например, томография. Подобные работы стали появляться еще в конце прошлого века. Например, в работе [12] использовался метод итеративного обратного проецирования, который, несмотря на высокую скорость сходимости, не позволяет получить уникальное решение, поскольку исходная задача является некорректной обратной задачей. Авторы алгоритма, представленного в [13], для нахождения единственного решения использовали метод регуляризации Тихонова [14]. Недостающие пиксели восстанавливались путем билинейной интерполяции, после чего их значения уточнялись в процессе итераций.

В настоящей работе использовалась более реалистичная, субпиксельная модель, в которой в формировании каждого пикселя низкого разрешения участвовали как центральные, так и априорно неизвестные доли соседних пикселей. Несмотря на то, что в предложенной модели появились новые параметры – сдвиги сетки из укрупненных пикселей относительно восстанавливаемых, имеющейся информации достаточно для восстановления даже величин этих сдвигов. Разработанный алгоритм, основанный на генеративно-состязательной нейронной сети, в качестве детектора которой использовалась математическая модель

формирования изображения низкого разрешения. Для выделения единственного решения использовался метод регуляризации с адаптивным выбором параметра регуляризации. Отметим, что разработанный алгоритм хорошо распараллеливается, поэтому может быть реализован аппаратно «на борту». Использование предложенного алгоритма в случае использования для съемки высокоскоростной камеры также позволяет увеличить ширину зоны охвата за счет изменения ориентации камеры – повороту ее таким образом, чтобы ее главная диагональ была перпендикулярна направлению движения летательного аппарата.

Метод повышения разрешения изображения

Пусть летательный аппарат, движущийся на постоянной высоте, проводит съемку местности с перекрытием кадров. Предположим, что из полученных снимков мы получили набор из K изображений X^1, \dots, X^K одного и того же участка местности. Пусть каждое из имеющихся изображений имеет размер $N \times M$ пикселей. Рассмотрим случай, когда требуется повысить разрешение в два раза, т.е. получить одно изображение размером $n \times t$ пикселей, где $n = 2N + 1$, $t = 2M + 1$. Визуально это выглядит так, как будто мы покрыли неизвестное изображение сеткой $n \times t$ пикселей, на которое наложили более крупную сетку, размер ячейки которой в два раза больше. Поскольку, вообще говоря, границы этих сеток в точности не совпадут, то в формировании каждого пикселя укрупненного изображения будет полностью участвовать один пиксель высокого разрешения и восемь частей соседних пикселей, как показано на рисунке 1.

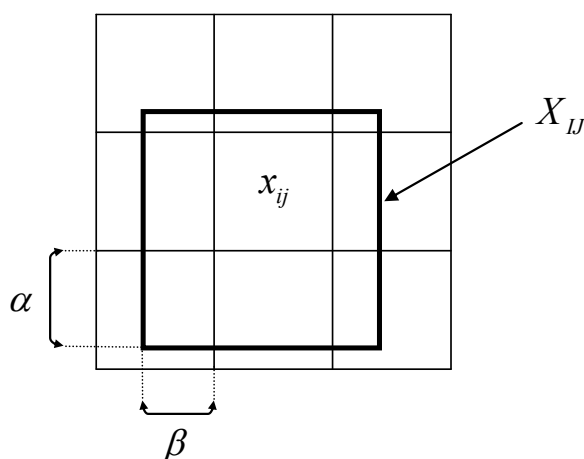


Рисунок 1 – Формирование пикселя низкого разрешения из 9 пикселей высокого разрешения

Обозначим интенсивность пикселя мелкой сетки в ячейке с номерами (i, j) через x_{ij} , тогда интенсивность пикселя с индексами (I, J) укрупненного k -го изображения X_{IJ}^k будет определяться по следующей формуле:

$$X_{IJ}^k = \alpha_k (\beta_k x_{i-1, j-1} + x_{i-1, j} + (1 - \beta_k) x_{i-1, j+1}) + (\beta_k x_{i, j-1} + x_{i, j} + (1 - \beta_k) x_{i, j+1}) + (1 - \alpha_k) (\beta_k x_{i+1, j-1} + x_{i+1, j} + (1 - \beta_k) x_{i+1, j+1}), \quad (1)$$

где $i = 2I, j = 2J$, а коэффициенты α_k и β_k характеризуют сдвиг крупной сетки относительно мелкой.

При этом должно выполняться условие

$$0 \leq \alpha_k \leq 1, 0 \leq \beta_k \leq 1. \quad (2)$$

Задача состоит в определении значений $x_{ij}, i = 1, \dots, n, j = 1, \dots, m$ и коэффициентов α_k и $\beta_k, k = 1, \dots, K$, по набору $X_{IJ}^k, I = 1, \dots, N, J = 1, \dots, M, k = 1, \dots, K$.

Для решения поставленной задачи использовалась нейронная сеть, принцип работы которой близок к так называемой генеративно-сопоставительной нейронной сети (GAN – Generative Adversarial Network [15]). Архитектура этой нейросети представлена на рисунке 2.

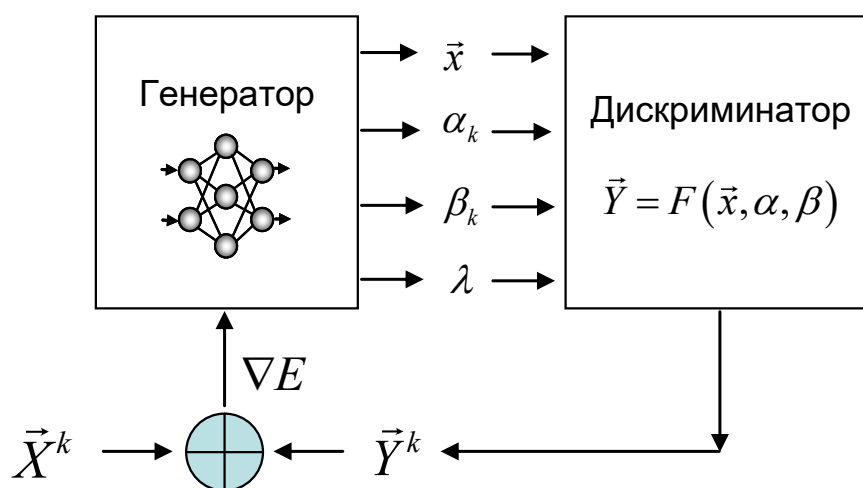


Рисунок 2 – Архитектура нейросети

Она состоит из двух частей: генератора и дискриминатора. Генератор в процессе настройки порождает как изображение высокого разрешения x_{ij} , так и другие параметры: величины сдвигов α_k, β_k и параметр регуляризации λ . Значения полученных величин подаются на

дискриминатор, который по формуле (1) получает изображение низкого разрешения Y^k . В результате сравнения со всеми имеющимися изображениями вычисляется ошибка генерации изображения:

$$E = \frac{1}{2} \sum_{k=1}^K \sum_{l=1}^N \sum_{j=1}^M (Y_{lj}^k - X_{lj}^k)^2, \quad (3)$$

градиент от которой ∇E используется при обучении нейронной сети генератора. Поскольку как значения параметра регуляризации λ , так и величины сдвигов α_k и β_k находятся в пределах сегмента $[0,1]$, то величины пикселей обрабатываемых изображений также были преобразованы к этому диапазону. За счет этого у выходных нейронов генератора была использована сигмоидная передаточная функция, что позволило для настройки нейросети избавиться от ограничений на величину весов. Отличием от широко используемых в настоящее время генеративно-сопоставительных нейронных сетей является то, что в качестве дискриминатора используется не нейронная сеть, а математическая модель рассматриваемого явления; в данном случае модель получения изображения низкого разрешения из изображения высокого разрешения. Этот принцип был предложен одним из авторов в 1999 году для реконструкции дефокусированных изображений [16].

Численный эксперимент

Поскольку для практического применения разработанного алгоритма наиболее важным параметром является его быстродействие, для настройки нейронной сети были протестированы различные современные оптимизационные методы: различные модификации Rprop (Resilient Propagation) [17], такие как Rprop+ [17], Rprop-, iRprop+, iRprop- [18], модификации метода Adam¹: Adagrad [19], Adadelta², Adamax³, а также хорошо известные методы сопряженных градиентов Флетчера-Ривса [20] и Полака-Райбера [21]. Особенностью алгоритмов семейств Rprop и Adam является их способность адаптировать скорость изменения каждого аргумента, что позволяет добиться существенного

¹ Kingma D.P. Adam: A Method for Stochastic Optimization // Cornell University Library. 2014. – URL: <https://arxiv.org/abs/1412.6980>

² Zeiler M. ADADELTA: An Adaptive Learning Rate Method // Cornell University Library. 2012. – URL: <https://arxiv.org/abs/1212.5701>

³ Kingma D.P. Op. cit.

ускорения процесса поиска минимума объектной функции по сравнению со стандартными методами типа BackProp [22], у которых используется одинаковый шаг оптимизации для всех переменных.

Для экспериментов использовались $L = 10$ различных спутниковых снимков высокого разрешения, загруженных с сайта <https://www.satimagingcorp.com>. Из них по формуле (1) получались изображения пониженного разрешения, которые использовались для обработки, а исходные изображения – для сравнения полученных результатов.

Процесс обучения нейросети был организован следующим образом. В начальный момент оно проводилось на одном изображении высокого разрешения, которое было получено из одного изображения низкого разрешения путем билинейной интерполяции. Далее, по достижении заданной точности, использовались все имеющиеся изображения низкого разрешения. Это позволило как ускорить процесс настройки, так и получать стабильно одинаковые конечные результаты вне зависимости от начальных значений коэффициентов α_k и β_k , которым назначались случайные значения в диапазоне $[\varepsilon, 1 - \varepsilon]$, где ε – малое число.

Качество обучения определялось как средняя ошибка в расчете на один пиксель, выраженная в процентах

$$E_C = \frac{100\%}{K \cdot N \cdot M} \sum_{k=1}^K \sum_{I=1}^N \sum_{J=1}^M |Y_{IJ}^k - X_{IJ}^k|. \quad (3)$$

Процесс обучения заканчивался при достижении E_C величины в 1%. После обработки всех выбранных изображений время обработки осреднялось. Самым быстрым оказался метод iRprop-. В таблице 1 представлены отношения времени работы наиболее быстрых методов по отношению ко времени выполнения iRprop-. В этой таблице аббревиатура CGFR (сокращенное наименование Conjugate Gradient Fletcher-Reeves) соответствует методу сопряженных градиентов в варианте Флетчера-Ривса.

Эти же характеристики отображены на рисунке 3.

Таблица 1 – Относительные скорости работы различных методов оптимизации для представленного алгоритма

iRprop-	Adadelta	RMSprop	iRprop+	CGFR	Rprop+	AdaMax	Rprop	Adam	Adagrad
1,00	1,04	1,17	1,79	1,83	2,67	3,25	3,75	3,92	3,92

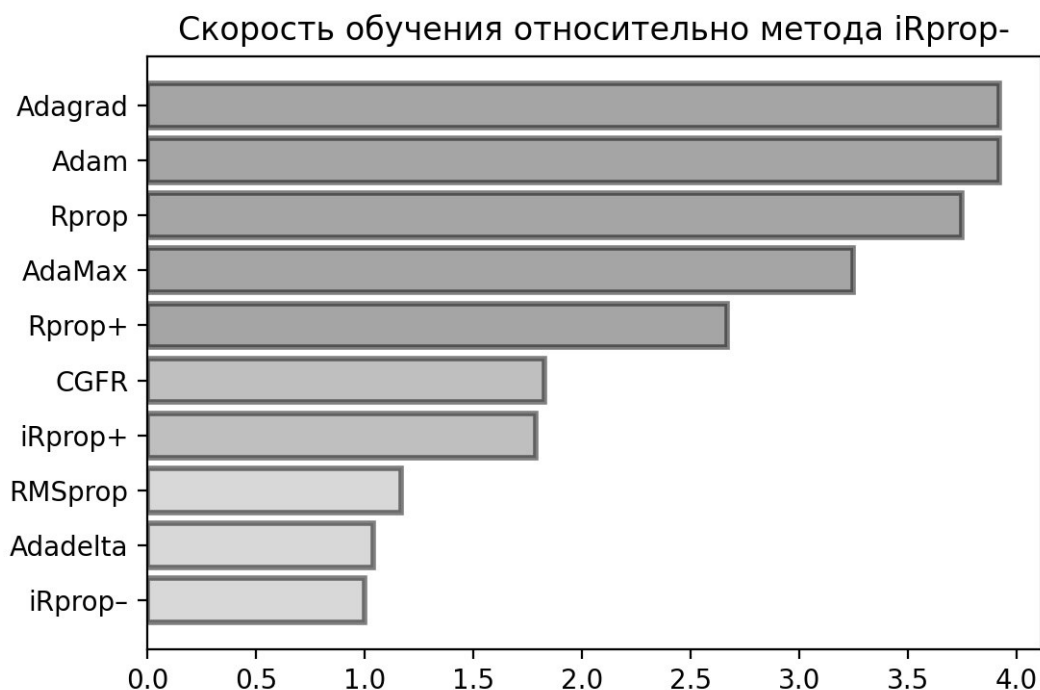


Рисунок 3 – Относительные скоростные характеристики обучения нейросети различными методами

Качество реконструкции изображения высокого разрешения определялось путем сравнения L полученных изображений высокого разрешения x_{ij}^l с исходными изображениями z_{ij}^l , по которым были получены изображения низкого разрешения. Напомним, что все изображения были приведены к диапазону $[0, 1]$, а в численном эксперименте участвовали $L = 10$ различных изображений. Для сравнения каждой l -й пары изображений использовалась следующая формула:

$$Q_l = \frac{100\%}{N_l \cdot M_l} \sum_{i=1}^{N_l} \sum_{j=1}^{M_l} |x_{ij}^l - z_{ij}^l|, \quad (4)$$

где $N_l \cdot M_l$ – число пикселей у l -го изображения, а потом показатель качества усреднялся по всем L изображениям, участвовавшим в эксперименте:

$$Q = \frac{1}{L} \sum_{l=1}^L Q_l.$$

Результаты сравнения приведены в таблице 2.

Таблица 2 – Точность Q реконструкции изображения высокого разрешения в процентах

Билинейная	Бикубическая	RAISR	Предложенный алгоритм; число изображений низкого разрешения K		
			$K=3$	$K=4$	$K=5$
7,94	7,64	7,14	3,73	3,59	2,42

Из таблицы видно, что разработанный алгоритм более чем в два раза превосходит другие алгоритмы по точности. На рисунке 4 приведены фрагменты одного и того же изображения, реконструированные по различным алгоритмам.



А – исходное изображение



Б – разработанный алгоритм



В – алгоритм RAISR



Г – Adobe Photoshop

Рисунок 4 – Фрагмент реконструированного изображения.
 А – исходное изображение, Б – восстановленное по разработанному алгоритму,
 В – алгоритмом RAISR, Г – бикубическая интерполяция

Повышение ширины зоны охвата

Обычно камера относительно летательного аппарата располагается таким образом, чтобы стороны фотоприемника были согласованы (параллельны) направлению его движения, как это показано на рисунке 5А. Нами предлагается расположить ее таким образом, чтобы диагональ фотоприемника была перпендикулярна направлению движения (рисунок 5Б); диагональ показана пунктиром.

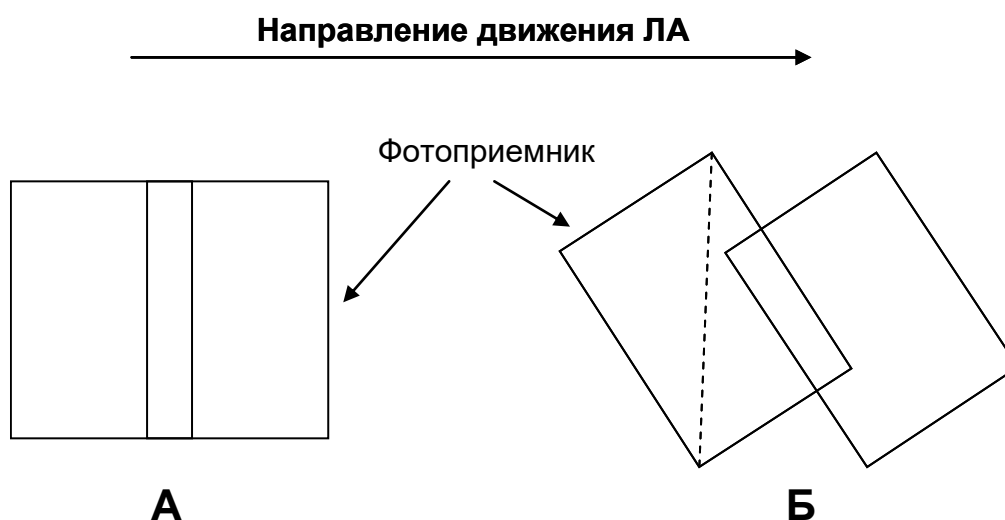


Рисунок 5 – А – стороны фотоприемника параллельны направлению движения ЛА, Б – главная диагональ фотоприемника перпендикулярна направлению движения ЛА

За счет этого ширина зоны охвата увеличится; например, если аспектное отношение сторон фотоприемника составляет 4:3, то ширина зоны увеличится на 25%. После этого для приведения изображения в «естественные» координаты, связанные с летательным аппаратом, нужно просто перенумеровать пиксели в соответствии со следующими формулами:

$$\begin{aligned}x' &= x \cos \gamma - y \sin \gamma, \\y' &= x \sin \gamma + y \cos \gamma,\end{aligned}$$

где γ – угол поворота диагонали фотоприемника. При таких условиях съемки всегда будет обеспечено смещение сеток на соседних кадрах, которое выше описывалось параметрами α_k и β_k .

Заключение

В настоящей работе предложен высокоэффективный алгоритм, позволяющий путем компьютерной обработки изображений, полученных в результате съемки с перекрытием изображений, повысить как разрешение, так и ширину зоны захвата оптико-электронной аппаратуры наблюдения до 25% при дистанционном наблюдении земной поверхности. Данный алгоритм отличается от существующих тем, что обладает высоким быстродействием, хорошо распараллеливается, требует минимальных вычислительных ресурсов и совместим с существующей отечественной элементной базой. Это позволит в дальнейшем реализовать его непосредственно в оптико-электронных системах наблюдения летательных аппаратов воздушного и космического базирования, а также осуществлять обработку оптико-электронных изображений на борту в реальном масштабе времени. Проведенные сравнения представленного алгоритма с существующими показали существенное, более чем в два раза, повышение точности реконструкции изображения высокого разрешения.

Список использованных источников

1. Бакланов А.И. К вопросу о пространственном разрешении и точности привязки изображений космических систем наблюдения высокого разрешения // Геоматика. 2010. Т.8. №3. – С. 25-30.
2. Архипов С.А. Исследование требований к перспективной оптико-электронной аппаратуре для малоразмерных космических аппаратов дистанционного зондирования Земли // Информация и космос. 2018. №1. – С. 155-162.
3. Rukundo O., Maharaj B.T. Optimization of image interpolation based on nearest neighbour algorithm // IEEE 2014 International Conference on Computer Vision Theory and Applications (VISAPP). 2014. Vol.1. – P. 641-647.
4. Gribbon K.T., Bailey D.G. A Novel Approach to Real-time Bilinear Interpolation // Proceedings of the Second IEEE International Workshop on Electronic Design, Test and Applications (DELTA 2004). 2004. – P. 126-131.
5. Hou H.S., Andrews H.C. Cubic splines for image interpolation and digital filtering // IEEE Trans. Acoustics, Speech, and Signal Processing. 1978. Vol.26. №6. – P. 508-517.
6. Johnson J., Alahi A., Li F. Perceptual losses for real-time style transfer and super-resolution // European conference on computer vision. 2016. – P. 18.
7. Shi W. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network / W. Shi, J. Caballero, F. Huszar, J. Totz, A.P. Aitken, R. Bishop, D. Rueckert, Z. Wang // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016. – P. 1874-1883.

8. Kim J., Kwon Lee J., Mu Lee K. Deeply-recursive convolutional network for image super-resolution // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016. – P. 1637-1645.
9. Ledig C., Theis L., Husza F. et al. Photo-realistic Single Image Super-Resolution Using a Generative Adversarial Network // Proc. IEEE Conf. Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. – P. 105-114.
10. Игнатъев В.Ю., Матвеев И.А., Мурынин А.Б., Усманова А.А., Цурков В.И. Повышение пространственного разрешения панхроматических спутниковых изображений на основе генеративных нейросетей // Известия РАН. Теория и системы управления. 2021. №2. – С. 62-70.
11. Romano Y., Isidoro J., Milanfar P. RAISR: Rapid and Accurate Image Super Resolution // IEEE Trans. on Comp. Imaging. 2017. Vol.3. №1. – P. 110-125.
12. Irani M., Peleg S. Super resolution from image sequences // Proceedings of International Conference on Pattern Recognition (ICPR'90). 1990. Vol.2. – P. 115-120.
13. Hong M.C., Kang M.G., Katsaggelos A.K. A regularized multichannel restoration approach for globally optimal high resolution video sequence // In SPIE VCIP, San Jose. 1997. Vol.3024. – P. 1306-1316.
14. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука, 1979. – С. 288.
15. Goodfellow I., Pouget-Abadie J., Mirza M. et al. Generative Adversarial Networks // Proc. Intern. Conf. Neural Information Processing Systems. Montreal, Canada. 2014. – P. 2672-2680.
16. Лукьяница А.А. Нейросеть для реконструкции дефокусированных изображений // Сб. научных трудов Всероссийской научно-технической конференции «Нейроинформатика-99». 1999. Ч.3. – С. 159-164.
17. Riedmiller M. Advanced supervised learning in multi-layer perceptrons – from backpropagation to adaptive learning algorithms // Computer Standards & Interfaces. 1994. V.16. №3. – P. 265-278.
18. Igel C., Hüsken M. Improving the Rprop Learning Algorithm // Proc. of Second International Symposium on Neural Computation (NC 2000). 2000. – P. 115-121.
19. Duchi J., Hazan E., Singer Y. Adaptive subgradient methods for online learning and stochastic optimization // Journal of Machine Learning Research. 2011. V.12. – P. 2121-2159.
20. Fletcher R. Practical methods of optimization. Wiley, 2000. – 450 p.
21. Гилл Ф., Мюррей У., Райт М. Практическая оптимизация. М.: Мир, 1985. – 509 с.
22. Rumelhart D.E., Hinton G.E., Williams R.J. Learning representations by back-propagating errors // Nature. 1986. Vol.323. – P. 533-536.