УДК 004.932

АДАПТИВНОЕ НЕЙРОСЕТЕВОЕ СЖАТИЕ МУЛЬТИСПЕКТРАЛЬНЫХ КОСМИЧЕСКИХ ИЗОБРАЖЕНИЙ ПОВЕРХНОСТИ ЗЕМЛИ

Ушенкин В.А., в.н.с. НИИ «Фотон» РГРТУ, к.т.н., foton@rsreu.ru

ADAPTIVE NEURAL NETWORK COMPRESSION OF MULTISPECTRAL SATELLITE IMAGES OF THE EARTH'S SURFACE

Ushenkin V.A.

The paper considers the problem of compressing multispectral images on board of remote sensing satellites in order to reduce the amount of information transmitted over the radio line. A simple architecture of artificial neural networks family, that allows adaptive compression of images with controlled distortions, is proposed. The question of training such family of networks is considered. The results of experimental studies are presented in comparison with known deterministic and neural network compression algorithms. In this paper, it was possible to obtain a two-stage neural network implementation of adaptive lossy image compression, which retains the advantages of known artificial neural networks over deterministic algorithms at high compression rates, but has significantly lower requirements for the speed and memory capacity of computing devices, which potentially allows it to be used on board of remote sensing satellites.

Key words: multispectral image, adaptive compression, remote sensing, artificial neural network.

Ключевые слова: мультиспектральное изображение, адаптивное сжатие, дистанционное зондирование Земли, искусственная нейронная сеть.

Введение

В последние годы с увеличением детальности съемки наблюдается стабильный рост объемов целевой информации, формируемой космическими аппаратами (КА) дистанционного зондирования Земли (ДЗЗ). Это создает про-

блемную ситуацию, связанную с обеспечением оперативности передачи на Землю большого по объему массива изображений при ограниченной пропускной способности радиолиний. Разрешение данной проблемы может быть осуществлено за счет бортового сжатия изображений перед их передачей по радиолинии.

Известен ряд классических алгоритмов, позволяющих уменьшить объем изображений. Алгоритмы сжатия без потерь (Deflate, RLE, LZW, LZMA, Lossless JPEG, Lossless JPEG2000) не вносят искажений в изображения, однако для космических снимков отличаются сравнительно невысокой степенью сжатия. Объем данных после применения указанных алгоритмов уменьшается примерно в 2 – 4 раза. В большей степени уменьшают объем данных алгоритмы сжатия с потерями (JPEG и JPEG2000), однако значительная степень сжатия достигается только при существенных искажениях изображения.

При передаче изображений с борта на Землю целесообразно прийти к компромиссу в виде сжатия с контролируемыми потерями, при которых основные характеристики изображения деградируют незначительно, зато при этом повышается оперативность передачи изображений, а также ослабляются ограничения на производительность КА. При контролируемых потерях с помощью алгоритмов JPEG и JPEG2000 можно уменьшить объем данных примерно в 5–30 раз.

Рассматривается задача сжатия мультиспектральных изображений на борту космических аппаратов дистанционного зондирования Земли с целью уменьшения объема информации, передаваемой по радиолинии. Предлагается простая архитектура семейства искусственных нейронных сетей, позволяющего адаптивно сжимать изображения с контролируемыми искажениями. Рассматривается вопрос обучения такого семейства сетей. Приводятся результаты экспериментальных исследований в сравнении с известными детерминированными и нейросетевыми алгоритмами сжатия.

> Также целесообразно применение адаптивного сжатия. На спутниковых снимках часто присутствуют малоинформативные объекты: облачность, водная поверхность и др. Фрагменты изображений с такими объектами можно сжимать с существенными потерями, дополнительно уменьшая объем передаваемых данных.

> Помимо классических детерминированных алгоритмов [1, 2], в последние годы опубликован ряд работ по применению искусственных нейронных сетей (ИНС) в задаче сжатия данных, в т.ч. изображений. Среди всех работ также можно выделить два крупных направления: сжатие изображений с потерями и сжатие данных без потерь. При этом подходы в рамках первого направления не исключают последующее применение подходов из второго направления к результатам сжатия с потерями, за счет чего степень сжатия увеличивается дополнительно. Результаты экспериментальных исследований, приводимые в опубликованных работах, показывают, что применение ИНС позволяет достичь при значительных степенях сжатия меньших искажений, чем в алгоритмах JPEG и JPEG2000. Однако известные нейросетевые подходы отличаются высокой вычислительной сложностью, что не позволяет их применять в условиях ограниченных вычислительных мощностей борта КА. Поэтому актуальна разработка нейросетевых подходов, обладающих существенно меньшей вычислительной сложно

стью, но сохраняющих преимущества своих прототипов. Такая разработка является целью настоящей работы.

Анализ известных нейросетевых подходов к сжатию изображений

В основе абсолютного большинства работ по нейросетевому сжатию изображений с потерями лежит использование так называемых автокодировщиков, представляющих собой ИНС, состоящую из двух частей: кодировщика и декодировщика [3-6]. При этом вход и выход ИНС идентичны по объему и структуре данных, а данные, передаваемые от кодировщика к декодировщику, имеют существенно меньший объем, чем данные на входе и выходе ИНС. При обучении автокодировщика добиваются наибольшего сходства данных на входе и выходе ИНС, при этом их отличие составляет потери при сжатии. Степень сжатия определяется структурой сети, а именно отношением объема данных на входе/выходе ИНС к объему данных, проходящих через центр сети. После обучения сеть разбивается на кодировщик и декодировщик, которые затем используются независимо, первый - при передаче данных, второй – при их приеме и распаковке.

Для управления степенью сжатия изображений с потерями используется два основных подхода. В первом подходе [3, 6] обучается несколько автокодировщиков, отличающихся объемом данных, проходящих через центр ИНС, а непосредственно при сжатии в соответствии с некоторым признаком для каждого сжимаемого фрагмента изображения адаптивно выбирается кодировщик с наиболее подходящей степенью сжатия. Во втором подходе [4, 5] используется рекуррентная схема. Обучается автокодировщик, характеризующийся большой степенью сжатия и значительными потерями. При выполнении сжатия изображение проходит через весь автокодировщик, после чего путем вычитания выходного и входного сигнала автокодировщика вычисляются потери. Сигнал, прошедший через центр автокодировщика, записывается в выходной поток. Если потери оказываются выше требуемых, то на вход автокодировщика подаются уже потери, результат сжатия которых записывается в выходной поток. Опять сравниваются сигналы на выходе и входе автокодировщика, и если они снова сильно отличаются друг от друга, рекуррентная схема сжатия продолжается.

Величина потерь при фиксированной степени сжатия существенно зависит от структуры автокодировщика. В опубликованных работах применяются многослойные персептроны [3], сверточные ИНС [6], рекуррентные ИНС [4, 5], слои нормализации данных [5, 6] и их комбинация. Многослойные персептроны не учитывают растровую структуру изображения и не позволяют добиться малых потерь. Рекуррентные ИНС хорошо подходят для рекуррентной схемы управления степенью сжатия, поскольку обладают памятью и по-разному преобразуют сигнал в зависимости от своего состояния. Однако они характеризуются наиболее высокой вычислительной сложностью. Поэтому для реализации на борту КА Д33 наиболее подходят сверточные ИНС, учитывающие растровую структуру изображения и обладающие умеренной вычислительной сложностью.

Результат сжатия с потерями может быть подвергнут дополнительному сжатию без потерь, которое может быть выполнено как с помощью традиционных алгоритмов, так и с помощью ИНС. В задаче сжатия без потерь ИНС используются для предсказания следующего значения сигнала по его предыдущим значениям. Результатом предсказания являются условные вероятности появления следующего символа после известных текущих. Переход от безусловных вероятностей появления символа в потоке к условным вероятностям позволяет существенно уменьшить энтропию сигнала, в результате чего последующее традиционное энтропийное кодирование существенно уменьшает объем данных. Результаты экспериментов, выполненных Мэттом Махони [7], показывают, что применение ИНС для предсказания следующего значения сигнала позволяет добиться значительно большего сжатия данных без потерь, чем в традиционных алгоритмах, ценой существенного увеличения времени сжатия и затрат оперативной памяти. Это недопустимо для бортовой реализации, однако возможно уменьшить затраты времени и памяти без существенного ухудшения степени сжатия. Основные затраты обусловлены необходимостью обучать ИНС при каждом сжатии новых данных. Их можно уменьшить, если удастся ИНС, используемую для предсказания, обучить заранее с обеспечением приемлемой точности предсказания для всех сжимаемых данных. Поскольку при бортовой обработке космических снимков сжатие без потерь применяется не к произвольным данным, а к результатам сжатия с потерями, всегда имеющим одну и ту же структуру и сходные зависимости между символами, такая задача может быть успешно решена.

Архитектура семейства автокодировщиков для адаптивного сжатия мультиспектральных изображений с потерями

Известные автокодировщики на базе сверточных ИНС имеют следующие общие черты. В них выполняется несколько последовательных сверток сигнала. Результаты сверток постепенно прореживаются по строкам и столбцам, а число каналов в изображении увеличивается. После прохождения нескольких сверточных слоев получается в несколько раз прореженное изображение с большим числом каналов. При этом в конце концов отношение текущего числа каналов к исходному оказывается меньше, чем квадрат коэффициента прореживания. За счет этого достигается уменьшение объема изображения. Для дополнительного уменьшения объема обычно выполняется квантование значений сигнала в пикселях сжатого изображения. Описанная часть ИНС представляет собой кодировщик. Декодировщик включает те же сверточные слои, расположенные в обратном порядке и чередующиеся с операцией «Depth to space». Указанная операция увеличивает число строк и столбцов изображения и уменьшает число каналов, заполняя новые строки данными отбрасываемых каналов. В результате восстанавливается изображение с исходным числом строк, столбцов и каналов. Выполнение в сверточных нейронных сетях нескольких последовательных

сверток с прореживанием позволяет построить пирамиду изображений, что имеет некоторое сходство с вейвлет-преобразованием, используемым в алгоритме JPEG2000 и учитывает растровую структуру изображения.

В алгоритме JPEG2000 используются одномерные фильтры из 7 и 9 элементов (вейвлеты Коэна-Добеши-Фово 9/7), применяемые сначала к строкам, а затем к столбцам изображения. Таким образом, преобразование фактически сводится к свертке в окне 9х9, которая описывается всего 9+7=16 коэффициентами. Свертка в пределах окна большей площади позволяет снизить влияние погрешностей квантования сигнала на верхнем уровне пирамиды на сигнал в отдельном пикселе на нижнем уровне пирамиды после восстановления сжатого изображения. Известные архитектуры автокодировщиков на каждом уровне пирамиды используют существенно большее число коэффициентов, но выполняют свертку с фильтрами меньшей апертуры (3х3 или 5х5). Для сокращения объема вычислений решено операцию «Conv2D» (свертки) заменить двумя операциями «DepthwiseConv2D» (поканальной свертки) в окнах 9х1 и 1x9 соответственно. Операция «DepthwiseConv2D» для формирования выходного канала изображения использует только один канал входного изображения, что существенно сокращает число весовых коэффициентов.

Кроме того, выполнен отказ от избыточного числа каналов на промежуточных уровнях пирамидального преобразования. На каждом последующем уровне число каналов увеличивается в 4 раза с одновременным прореживанием по строкам и столбцам в 2 раза и сохранением общего объема данных. По окончании пирамидального преобразования на верхнем уровне пирамиды выполняется дополнительная свертка сигнала с уменьшением числа каналов и последующим квантованием с целью уменьшения объема данных.

Поскольку при сжатии с потерями в первую очередь деградирует высокочастотная составляющая сигнала изображения, решено добавить в архитектуру автокодировщика слой коррекции этой составляющей. Он заключается в поканальной свертке сигнала в окне 5х5 с фильтром, сумма коэффициентов которого равна нулю, и прибавлении результата свертки к исходному сигналу.

Перед кодированием выполняется нормализация яркостей изображения, а в конце декодирования – денормализация. Также в начале и в конце автокодировщика выполняется операция «Conv2D» в окне 1x1, по сути выполняющая некоторое преобразование цветового пространства изображения по аналогии с алгоритмами JPEG и JPEG2000. Полученная архитектура автокодировщика приведена на рис. 1.

Недостатком известных адаптивных подходов к сжатию является возникновение блочных артефактов при переключении степени сжатия. Это связано с тем, что для каждой степени сжатия автокодировщики работают с разными весовыми коэффициентами во всех слоях. Чтобы избежать влияния краевых эффектов на свертки, выполняемые при декодировании сжатого изображения, требуется записывать в сжатый поток избыточные данные, что отрицательно сказывается на степени сжатия. В противном случае на границах блоков возникают краевые эффекты, увеличивающие уровень искажений при сжатии и хорошо заметные глазу.

Для устранения этого эффекта предлагается принципиально новый подход к обучению ИНС для сжатия: обучение семейства взаимосогласованных по весовым коэффициентам кодировщиков для каждой степени сжатия с единым декодировщиком.

Все кодировщики в рамках семейства имеют одни и те же весовые коэффициенты для пирамидального преобразования изображения и отличаются друг от друга лишь на последнем уровне. Последние уровни кодировщиков формируют разное число X выходных каналов. При этом весовые коэффициенты кодировщика с большей степенью сжатия, формирующего меньшее число выходных каналов, целиком включаются в состав выходного слоя кодировщика с меньшей степенью сжатия, формирующего те же выходные каналы, а также ряд дополнительных выходных каналов.

количество квантованных каналов сжатого изображения. Единый декодировщик позволяет из всех сжатых блоков собрать единое сжатое изображение. Если при некоторой степени сжатия в блоке формируется меньшее число каналов, недостающие каналы заполняются нулями. В этом также прослеживается сходство с алгоритмами JPEG/JPEG2000, в которых обнуляется часть коэффициентов. Единое сжатое изображение преобразуется декодировщиком в результат распаковки, на котором отсутствуют блочные артефакты, поскольку резкая граница между нулевым и ненулевым сигналом соседних блоков размывается в ходе сверток, выполняемых декодировщиком.

Основные трудности возникают с обучением такого семейства нейронных сетей. Добиться минимума потерь на всех степенях сжатия крайне затруднительно, поскольку для этого требуется многокритериальная оптимизация по нескольким выходам нейронной сети. В связи с этим предлагается «жадный» алгоритм обучения.

Сначала подбираются коэффициенты пирамидального преобразования изображения, обеспечивающие минимальное влияние шумов квантования сигнала верхнего уровня пирамиды на восстановление изображения на нижнем уровне. Для этого строится базовый автокодировщик, не выполняющий сжатие изображения за счет уменьшения числа каналов. Сжатие в нем выполняется только лишь за счет квантования сигнала верхнего уровня пирамиды с минимальным числом бит. В рамках настоящей работы в качестве такого числа бит выбрано 4.

Для обучения семейства кодировщиков последовательно составляются и обучаются автокодировщики с разным числом выходных каналов в слое перед блочноадаптивным квантованием. Первым обучается автокодировщик с максимальной степенью сжатия (минимальным числом каналов, равным 12, каждый из которых адаптивно квантуется 4 битами на отрезке от минимального до максимального значения сигнала в пределах блока). Границы адаптивного квантования в пределах блока вместе с собственно квантованными значениями записываются в выходной поток.





Рис. 1. Предлагаемая архитектура автокодировщика

При обучении такого автокодировщика в начальные и конечные слои сети загружаются и фиксируются весовые коэффициенты, полученные для автокодировщика, обученного ранее, а обучение выполняется только для двух средних слоев, соседних со слоем квантования.

Декодировщик формируется единый для всех кодировщиков и принимает на вход максимальное

Для следующего автокодировщика увеличивается до 24 число каналов на выходе слоя перед квантованием и на входе слоя после квантования. Каждый канал попрежнему квантуется 4 битами. В слои данного автокодировщика загружаются и фиксируются весовые коэффициенты от предыдущей сети. В средних слоях новой сети число весовых коэффициентов увеличивается. Поэтому в них загружается и фиксируется лишь часть весов, а оставшаяся часть подбирается в ходе обучения. Таким образом, новый кодировщик формирует те же 12 каналов сжатого изображения, что и предыдущий, а также 12 новых каналов, позволяющих более точно восстановить исходное изображение из сжатого.

Для следующего автокодировщика число каналов составляет 36. Первые 24 канала квантуются 5 битами и формируются с теми же весами, что и в предыдущей сети, а дополнительные 12 каналов формируются с весами, подбираемыми в ходе обучения, и квантуются 6 битами.

Для последнего автокодировщика число каналов составляет 60. Первые 36 каналов квантуются 6 битами, а дополнительные 24 канала – 7 битами. Таким образом, сигнал от каждого следующего кодировщика целиком включает в себя сигналы от всех предыдущих кодировщиков, лишь менее грубо их квантуя и дополняя новыми каналами, содержащими поправки для более точного восстановления исходного изображения. Вторая половина слоев последнего автокодировщика образует единый декодировщик адаптивно сжатого изображения.

Объемы данных на выходе кодировщиков составляют 0,75, 1,5, 3 и 6 бит на пиксель исходного мультиспектрального 3-канального изображения.

Для адаптивного выбора степени сжатия с приемлемым уровнем искажений может использоваться следующий подход. В рамках этого подхода сигнал проходит через последний обученный кодировщик, выполняющий минимальное сжатие. Затем выполняется отбрасывание (обнуление) части каналов на выходе кодировщика и квантование сигнала в оставшихся каналах. Сначала оставляются и квантуются 4 битами 12 каналов, соответствующих максимальному сжатию. Полученный сигнал подается на вход декодировщика, сигнал на выходе которого сопоставляется с блоком исходного изображения с вычислением потерь при сжатии. Если потери слишком велики, то от сигнала, сформированного кодировщиком, оставляется большее количество каналов и вновь выполняется декодирование с оценкой потерь. Процесс повторяется либо до достижения при некоторой степени сжатия приемлемых потерь, либо до достижения минимально предусмотренной степени сжатия. Эксперименты показали, что среднеквадратичная ошибка восстановления исходного изображения по сжатому для каждого следующего автокодировщика уменьшается примерно в 1,4 раза по сравнению с предыдущим.

В дополнение к описанному процессу может выполняться учет дополнительных признаков. Если имеется маска малоинформативных объектов на снимке, а сигнал сжимаемого блока целиком соответствует малоинформативным объектам, для этого блока выбирается максимальная степень сжатия без контроля потерь.

Архитектура семейства нейросетевых предсказателей для дополнительного сжатия изображений без потерь

Результаты нейросетевого сжатия с потерями могут быть подвергнуты дополнительному сжатию без потерь на основе энтропийного кодирования и нейросетевого предсказания значения следующего отсчета сигнала по предыдущим значениям. Такой подход позволяет добиться большего сжатия без потерь результатов нейросетевого сжатия с потерями, чем детерминированные алгоритмы. Экспериментальные исследования показали, что традиционные архиваторы WinRAR и 7 Zip уменьшают объем результатов нейросетевого сжатия с потерями всего на ~5 %, а лучшие из имеющихся детерминированных алгоритмов (МСМ и BSC) – на 10-11 %. В то же время с помощью открытого алгоритма нейросетевого сжатия без потерь DeepZIP [8] удалось уменьшить объем на ~20 %.

При этом был поставлен эксперимент, в рамках которого ИНС-предсказатель в DeepZIP была обучена по одной пятой части изображения, а затем применена ко всему изображению. Достигнутая степень сжатия оказалась сопоставима как для пятой части файла, по которой выполнялось обучение, так и для файла целиком. Это позволило сделать вывод, что нейронную сетьпредсказатель достаточно обучить один раз, а затем многократно использовать для сжатия без потерь результатов сжатия изображений с потерями.

В составе нейронных сетей-предсказателей в настоящее время используются следующие слои: рекуррентные слои (LSTM, GRU) и полносвязные слои. По результатам экспериментальной отработки с помощью алгоритма DeepZIP, поддерживающего несколько различных архитектур сети-предсказателя, установлено, что наилучшее предсказание следующего пикселя результатов сжатия с потерями обеспечивает нейронная сеть на базе полносвязных слоев. Достигнутый объем данных получается на 1 % ниже, чем при использовании слоев LSTM и GRU, при этом объем вычислений оказывается в 2 раза меньше.

Следующим шагом по повышению степени сжатия без потерь стало обучение отдельной сети на предсказание следующего пикселя в каждом канале изображения, сжатого нейронной сетью с потерями. Нейронная сеть использует значения каналов предыдущего пикселя изображения и значения предыдущих каналов текущего пикселя, чтобы предсказать значение следующего канала текущего пикселя. Число нейронных сетейпредсказателей равняется числу каналов. В результате с помощью набора нейронных сетей на базе полносвязных слоев, каждая из которых была обучена по отдельному каналу изображения, достигнуто на несколько процентов обльшее уменьшение объема данных, чем у DeepZIP. При этом в каждом отдельном предсказателе удалось уменьшить число нейронов без существенного снижения достигаемой степени сжатия. В итоге, наилучнаилучшим образом учтя структуру сжимаемых данных, удалось добиться высокой степени сжатия без потерь при небольших вычислительных затратах и умеренном потреблении оперативной памяти. Достигнутая степень сжатия лишь немногим уступает лучшему на сегодняшний день алгоритму сжатия без потерь NNCP [9] на основе ИНС-трансформера, требующему на несколько порядков больших затрат.

Архитектура семейства поканальных предсказателей на основе полносвязных ИНС представлена на рис. 2, где *b* – число бит, которыми квантуется сигнал на выходе сжатия с потерями. Размерность сигнала, подаваемого на вход предсказателя, равняется числу каналов при заданной степени сжатия.



Рис. 2. Предлагаемая архитектура предсказателя

Обучение ИНС

Для обучения автокодировщиков необходимы разрезанные на фрагменты мультиспектральные космические снимки с пространственным разрешением, соответствующим характеристикам заданного КА ДЗЗ. Снимки должны содержать большинство типовых наблюдаемых объектов в пределах всего возможного диапазона яркостей: водную поверхность, облака, поля, леса, городскую застройку. Расчет целевой функции обучения выполняется лишь по центральной части каждого фрагмента для исключения влияния краевых эффектов. В качестве целевой функции могут использоваться среднеквадратичное отклонение яркостей изображений на входе и выходе автокодировщика, одномасштабный и многомасштабный индексы структурного сходства и другие показатели. В настоящей работе при обучении использовалось среднеквадратичное отклонение.

Из-за наличия слоя квантования, отличающегося разрывным градиентом, сходимость процесса минимизации целевой функции крайне нестабильная. При этом применение пакетной нормализации [10] в данном случае сходимость практически не улучшает, поэтому ее использование нецелесообразно. За ходом обучения необходимо постоянно следить и снижать скорость градиентного спуска, если процесс начал расходиться или целевая функция перестала уменьшаться. Особенно сложно выполняется обучение базового автокодировщика, поскольку подбираются коэффициенты для всех слоев нейронной сети (так называемое глубокое обучение).

Для обучения нейронных сетей-предсказателей генерируются отдельные выборки путем неадаптивного сжатия с потерями мультиспектральных космических снимков. Снимки должны быть сжаты с каждой из возможных степеней сжатия, фиксированной в пределах всего снимка. На вход каждого предсказателя значения в k-м канале сжатых данных (k изменяется от 1 до Nк) подаются значения в последних (Nк-k+1) каналах предыдущего пикселя сжатого изображения и в первых (k-1) каналах текущего пикселя. Эталонным выходом нейронной сети является сигнал из 2b отсчетов, из которых один, соответствующий b-битному коду в k-м канале текущего пикселя, является единичным, а остальные нулевыми. Целевой функцией, минимизируемой при обучении, является перекрестная энтропия между выходом нейронной сети и эталонным выходом. В отличие от обучения автокодировщиков, включающего от нескольких сотен до нескольких тысяч эпох, обучение каждого предсказателя выполняется быстро: за несколько эпох.

Экспериментальные исследования

Проведены экспериментальные исследования по сравнению потерь при реализованном нейросетевом сжатии по сравнению с детерминированными алгоритмами сжатия изображений с потерями JPEG и JPEG2000, а также с одним из известных нейросетевых алгоритмов [4], для которого в открытом доступе имеется обученная рекуррентная ИНС. Сравнение с другими нейросетевыми алгоритмами сжатия с потерями затруднительно ввиду отсутствия обученных ИНС.

В качестве меры потерь при сжатии использовались следующие показатели:

 – среднеквадратичное отклонение яркостей в каналах пикселей исходного и распакованного сжатого изображений;

 индекс структурного сходства и его три составляющие: сходство яркостей, сходство контраста и сходство структуры [11];

 интегральный показатель сходства яркостей, контраста и структуры.

Среднеквадратичное отклонение рассчитывается как

$$\sigma=\sqrt{rac{1}{N}\sum\limits_{i=1}^{N}ig(b_{ucxi}-b_{ ext{cxi}}ig)^2}$$
 , где b_{ucxi} – код яркости i -го от-

счета исходного изображения, $b_{\rm CMCI}$ – код яркости *i*-го отсчета распакованного сжатого изображения, N – общее число отсчетов, равное произведению числа пикселей и числа каналов. За счет большого числа отсчетов смещением на единицу числа степеней свободы при оценке выборочной дисперсии пренебрегается. Отсутствию искажений соответствует нулевое значение среднеквадратичного отклонения.

Индекс структурного сходства рассчитывается для каждого канала в скользящем окне 8х8 пикселей, а в качестве итогового значения берется результат усреднения по всем положениям скользящего окна во всех каналах изображений. В пределах каждого окна вычисляются статистические характеристики яркостей в выбранном канале исходного и сжатого изображений: выборочные средние и дисперсии, а также ковариация между яркостями исходного и сжатого изображений. Индекс структурного сходства в пределах окна рассчитывается как SSIM =

$$=\frac{(2\mu_{\text{HCX}}\mu_{\text{CX}}+(0,01\cdot255)^2)(2\operatorname{cov}_{\text{HCX},\text{CX}}+(0,03\cdot255)^2)}{(\mu_{\text{HCX}}^2+\mu_{\text{CX}}^2+(0,01\cdot255)^2)(\sigma_{\text{HCX}}^2+\sigma_{\text{CX}}^2+(0,03\cdot255)^2)},$$

где $\mu_{\rm Hex}$ и $\sigma_{\rm Hex}^2$ – выборочные среднее и дисперсия яркостей в канале исходного изображения, $\mu_{\rm cw}$ и $\sigma_{\rm cw}^2$ – выборочные среднее и дисперсия яркостей в канале распакованного сжатого изображения, со $v_{\rm Hex,cw}$ – ковариация между яркостями исходного и сжатого изображений.

Значение индекса структурного сходства в пределах каждого окна для каждого канала может быть представлено в виде произведения трех составляющих: сходства

яркостей
$$l = \frac{2\mu_{\text{нск}}\mu_{\text{ск}} + (0,01 \cdot 255)^2}{\mu_{\text{нск}}^2 + \mu_{\text{ск}}^2 + (0,01 \cdot 255)^2}$$
, сходства контра-

ста $c = \frac{2\sigma_{\text{исх}}\sigma_{\text{сж}} + (0,03 \cdot 255)^2}{\sigma_{\text{исx}}^2 + \sigma_{\text{сж}}^2 + (0,03 \cdot 255)^2}$ и сходства структуры

 $s = rac{\mathrm{cov}_{_{\mathrm{HCX,CK}}} + 0,5 \cdot (0,03 \cdot 255)^2}{\sigma_{_{\mathrm{HCX}}}\sigma_{_{\mathrm{CK}}} + 0,5 \cdot (0,03 \cdot 255)^2}.$ Последний показатель

представляет собой коэффициент корреляции Пирсона, в числитель и знаменатель которого добавлены регуляризующие коэффициенты, позволяющие избежать деления на нуль. Перечисленные составляющие также усредняются по всем положениям скользящего окна и всем каналам с получением трех частных показателей сходства изображений. Индекс структурного сходства и три его составляющие при отсутствии искажений принимают единичное значение, а с увеличением искажений при сжатии снижаются. Поэтому в качестве меры потерь могут использоваться величины 1-SSIM, 1-l, 1-c и 1-s.

Индекс структурного сходства является широко распространенной мерой оценки качества нейросетевого сжатия, однако разложение его на составляющие показывает его недостатки. Показатель сходства яркостей сопоставляет средние яркости, но при наличии искажений оказывается не инвариантен к одинаковому аддитивному сдвигу яркостей сравниваемых изображений: $l(a,b) \neq l(a+c,b+c)$. Если средние яркости оказываются малы, то их небольшие различия начинают сильно снижать значение показателя, в то время как такие искажения практически не заметны человеческому глазу и практически не ухудшают измерительные свойства изображения. Частично этот недостаток компенсирует наличие регуляризующего коэффициента в числителе и знаменателе показателя, однако негативное влияние всё равно присутствует. Аналогичные рассуждения можно провести и для показателя сходства контрастов. В случае сопоставления изображений однородных по яркости объектов, когда $\sigma_{_{\rm HCX}}^2$ и $\sigma_{_{\rm cc}}^2$ малы, значение показателя получается заниженным.

С учетом вышесказанного в рамках настоящей работы предлагается новый интегральный показатель сходства яркостей, контраста и структуры, выступающий альтернативой индексу структурного сходства. В пределах каждого положения скользящего окна в каждом канале он

рассчитывается как
$$LCS = \frac{\text{cov}_{\text{исх,ск}} + 0.5 \cdot (0.03 \cdot 255)^2}{\sigma_{\text{исх}} \sigma_{\text{ск}} + 0.5 \cdot (0.03 \cdot 255)^2}$$

$$-\frac{|\mu_{_{\rm HCX}}-\mu_{_{\rm CK}}|}{255}-\frac{6\cdot|\sigma_{_{
m HCX}}-\sigma_{_{
m CK}}|}{255}$$
. Коэффициент 6 предпола-

гает, что яркости наиболее контрастных изображений имеют нормальное распределение на отрезке [0, 255].

При выполнении экспериментальных исследований выполнялось сжатие изображений алгоритмом JPEG с разными значениями параметра качества, алгоритмом JPEG2000 с разными значениями степени сжатия и предложенным адаптивным нейросетевым алгоритмом с разными значениями допустимых потерь в двух конфигурациях: без дополнительного сжатия без потерь и с дополнительным сжатием без потерь на основе нейросетевого предсказания и энтропийного кодирования. Графики полученных значений показателей приведены на рис. 3-8.













Из рисунков видно, что нейросетевое сжатие мультиспектральных изображений имеет преимущество над детерминированными алгоритмами по части показателей при высоких степенях сжатия. При этом преимущество достигается только в сочетании с последующим сжатием без потерь на основе нейросетевого предсказания. При степенях сжатия, больших 28, нейросетевое сжатие превосходит алгоритм JPEG по критерию среднеквадратичного отклонения яркостей, но всё равно уступает по данному критерию алгоритму JPEG2000. Нейросетевое сжатие лучше показывает себя по индексу структурного сходства. Начиная со степени сжатия 25, оно существенно превосходит алгоритм JPEG, а со степени сжатия 38 начинает незначительно превосходить алгоритм JPEG2000. Наибольший выигрыш нейросетевого сжатия наблюдается по введенному интегральному показателю сходства яркостей, контраста и структуры. Оно существенно превосходит алгоритм JPEG, начиная со степени сжатия 29, и примерно на 1 % превосходит алгоритм JPEG2000, начиная со степени сжатия 33.



Рис. 6. Зависимость сходства контраста от степени сжатия мультиспектрального изображения







Рис. 8. Зависимость интегрального показателя сходства яркостей, контраста и структуры от степени сжатия мультиспектрального изображения

Частные показатели показывают лишь отдельные аспекты сходства исходного и сжатого изображений и позволяют выявить, какого типа искажения в основном вносит тот или иной алгоритм сжатия. Слабым местом алгоритма JPEG2000 является искажение локального контраста изображения. Слабым местом ИНС – искажение среднего уровня яркости. Слабым местом алгоритма JPEG – искажение структуры изображения.

Скорость предлагаемого нейросетевого сжатия на процессоре Intel Core i7-6800К составила около 0,32 Мпикс/с, что в ~12 раз медленнее алгоритма JPEG2000, но уже потенциально реализуемо на борту КА Д33.

Сравнение с рекуррентной ИНС [4] выполнено на небольшом фрагменте мультиспектрального спутникового снимка городской территории (размером 4000х2000 пикселей) из-за огромной вычислительной сложности обработки с помощью этой ИНС. Время обработки указанного фрагмента рекуррентной нейронной сетью на процессоре Intel Core i7-6800К составило около 3 часов, а затраты оперативной памяти – около 45 Гбайт. В то же время обработка с помощью предложенной в настоящей работе архитектуры заняла 18 секунд при затратах оперативной памяти в десятки Мбайт. Поскольку для рекуррентной ИНС [4] в открытом доступе имеется полноценная реализация лишь первой части сжатия (сжатие с потерями на базе автокодировщика), а для последующего сжатия без потерь реализация отсутствует, при сравнении с ней предлагаемая архитектура автокодировщиков также использовалась без последующего дополнительного сжатия без потерь. Результаты сравнения приведены на рис. 9-14. Из рисунков видно, что нейронные сети очень близки по значениям основных показателей. Рекуррентная ИНС [4] лучше по среднеквадратичному отклонению и сходству яркостей, а предложенная сеть лучше по индексу структурного сходства и сходству контрастов. По сходству структуры и введенному интегральному показателю сети практически одинаковы.

Заключение

В настоящей работе удалось получить двухэтапную нейросетевую реализацию адаптивного сжатия изображений с потерями, которая сохраняет преимущества известных ИНС перед детерминированными алгоритмами при высоких степенях сжатия, но обладает существенно меньшими требованиями к быстродействию и объему памяти вычислительных устройств, что потенциально позволяет ее использовать на борту КА ДЗЗ.



Рис. 9. Зависимость среднеквадратичного отклонения от степени сжатия тестового фрагмента мультиспектрального изображения







Рис. 11. Зависимость сходства яркостей от степени сжатия тестового фрагмента мультиспектрального изображения



Рис. 13. Зависимость сходства структуры от степени сжатия тестового фрагмента мультиспектрального изображения

Литература

1. Косткин И.В. Алгоритм вейвлет-сжатия неподвижных цифровых изображений. Вестник Рязанского государственного радиотехнического университета. 2007. № 20. С. 114-117.

2. Васильев А.М. Алгоритм сжатия изображений на основе модифицированного дискретного косинусного преобразования. Вестник Рязанского государственного радиотехнического университета. 2013. № 46-2. С. 16-21.

3. Sahoolizadeh H., Suratgar A.A. Adaptive image compression using neural networks. 5th International Conference: Science of Electronics, Technologies of Information and Telecommunications. 2009.

4. Toderici G., Vincent D., Johnston N., Hwang S.J., Minnen D., Shor J., Covell M. Full resolution image compression with recurrent neural networks. IEEE Conference on Computer Vision and Pattern Recognition. 2017.

5. Islam K., Dang M., Lee S., Moon H. Image compression with recurrent neural networks and generalized divisive normalization. IEEE Conference on Computer Vision and







Рис. 14. Зависимость интегрального показателя сходства яркостей, контраста и структуры от степени сжатия тестового фрагмента мультиспектрального изображения

Pattern Recognition. 2021.

6. Theis L., Shi W., Cunningham A., Huszar F. Lossy image compression with compressive autoencoders. International Conference on Learning Representations. 2017.

7. Mahoney M. Large Text Compression Benchmark. URL: http://www.mattmahoney.net/dc/text.html (дата обращения 01.02.2023).

8. Goyal M., Tatwawadi K., Chandak S., Ochoa I. DeepZip: Lossless data compression using recurrent neural networks. Data Compression Conference. 2019.

9. Bellard F. Lossless Data Compression with Neural Networks. URL: https://bellard.org/nncp/nncp.pdf (дата обращения 01.02.2023).

10. loffe S., Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. Proceedings of the 32nd International Conference on International Conference on Machine Learning. 2015, vol. 37, pp. 448-456.

11. Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P. Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing. 2004.