

СТАТЬИ

УДК 004.042

**РАЗРАБОТКА И ИССЛЕДОВАНИЕ МЕТОДОВ СЖАТИЯ ДАННЫХ
ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ**

Абрамов Н.С., Шишкин О.Г.

*ФГБУН Институт программных систем им. А.К. Айламазяна Российской академии наук,
Веськово, e-mail: psi@botik.ru*

В системах дистанционного зондирования Земли (ДЗЗ) непрерывный рост размера данных усложняет обработку информации, из-за этого возникает проблема процесса ускорения, которая решается улучшением способов компрессии информации. Для задач изучения Земли из космоса составляет интерес компрессии информации с небольшими потерями или без потерь. Основываясь на обзор по сжатию существует множество методов, схем или алгоритмов сжатия, в которых присутствуют свои недоработки или достоинства. В связи с системным ростом данных ДЗЗ, в том числе проблема хранения информации задача компрессии, на сегодняшний день, является актуальной. Решение данной проблемы с помощью методов сжатия обеспечивает: повышение эффективности управления сложными техническими системами за счет оперативного получения важной информации из сжатых данных; уменьшение больших потоков цифровых данных; повышение памяти выходных устройств записи информации; повышение полосы пропускания канала. В статье проведен аналитический обзор работ по компрессии информации для систем ДЗЗ. Представлены метод и результаты сжатия мультиспектральных изображений с использованием архиватора PAQ для повышения пропускной способности в системах дистанционного зондирования Земли.

Ключевые слова: PAQ, сжатие информации, дистанционное зондирование Земли, мультиспектральные снимки

**DEVELOPMENT AND RESEARCH OF DATA COMPRESSION
METHODS OF REMOTE EARTH SENSING**

Abramov N.S., Shishkin O.G.

*Ailamazyan Program Systems Institute of Russian Academy of Sciences,
Veskovo, e-mail: psi@botik.ru*

In Earth remote sensing systems (ERS), the continuous growth of the data size complicates the processing of information, because of this, the problem of the acceleration process arises, which is solved by improving the methods of information compression. For the tasks of studying the Earth from space, it is of interest to compress information with little or no loss. Based on the review on compression, there are many compression methods, schemes or algorithms that have their own drawbacks or advantages. In connection with the systemic growth of remote sensing data, including the problem of storing information, the task of compressing, today, is urgent. The solution to this problem with the help of compression methods provides: increasing the efficiency of managing complex technical systems due to the prompt receipt of important information from compressed data; reduction of large streams of digital data; increasing the memory of the output devices for recording information; increasing the channel bandwidth. The article provides an analytical review of work on data compression for remote sensing systems. The method and results of compression of multispectral images using the PAQ archiver to increase the throughput in Earth remote sensing systems are presented.

Keywords: PAQ, information compression, Earth remote sensing, multispectral imagery

В системах ДЗЗ постоянный рост объема информации затрудняет обработку данных, при которых появляется сложность процесса ускорения, которая решается усовершенствованием методов сжатия данных. Решение данной проблемы с помощью методов сжатия обеспечивает: повышение эффективности управления сложными техническими системами за счет оперативного получения важной информации из сжатых данных; уменьшение больших потоков цифровых данных; повышение памяти выходных устройств записи информации; повышение полосы пропускания канала.

Актуальность задачи компрессии связана с корректным хранением информации ДЗЗ, особенно растровые изображения используют существенный размер памяти.

Важность проблемы сжатия в том числе связана с проблемой корректного хране-

ния данных дистанционного зондирования, поскольку графические данные, особенно файлы растровых изображений, занимают значительный объем памяти. Решение данной проблемы с помощью методов сжатия обеспечивает: повышение эффективности управления сложными техническими системами за счет оперативного получения важной информации из сжатых данных; уменьшение больших потоков цифровых данных; повышение памяти выходных устройств записи информации; повышение полосы пропускания канала. В статье проведен аналитический обзор работ по компрессии информации для систем ДЗЗ.

Цель исследования: проведение аналитического обзора работ по сжатию данных для систем ДЗЗ. Тестирование метода сжатия мультиспектральных изображений с использованием архиватора PAQ с целью

повышения эффективности их передачи, хранения и дальнейшего анализа.

Материалы и методы исследования

1. Методы сжатия данных ДЗЗ

Достаточно много имеется разных типов алгоритмов компрессии или декомпрессии с разными характеристиками, которые используются в различных областях. Анализ существующих методов показал, что аппарат искусственных нейронных сетей (ИНС) активно используется для решения задач сжатия целевых данных ДЗЗ. Рассмотрим некоторые работы в этой области. В статье [1] представлена технология сжатия многоспектральных космических снимков. Рассмотрена возможность применения преобразования главных компонент для декорреляции массива многоспектральных данных. Предложен алгоритм сжатия главных компонент методом Хаффмана с использованием масштабных коэффициентов.

В статье [2] предлагаются алгоритмы для сжатия изображений без потерь и с потерями, в том числе учитывающие вероятность повторения значений пикселей и, таким образом, обеспечивающие увеличение степени сжатия. Авторы используют метод гистограмм и алгоритм RLE, который хоть и не отличается высоким коэффициентом сжатия, но имеет крайне низкую вычислительную сложность. Другой метод сжатия изображений без потерь, требующий минимальных вычислительных ресурсов, предложен авторами статьи [3]. Метод не имеет спектральных преобразований и устраняет статистическую избыточность данных ДЗЗ.

Ежедневно огромное количество информации хранится, обрабатывается и передается в цифровом виде по всему миру. Нейронные сети были быстро разработаны и исследованы как решение задач обработки изображений и контроля исправления ошибок канала. В работе [4] используется глубокая нейронная сеть (DNN) для сжатия серого изображения и отказоустойчивую систему передачи с возможностями коррекции ошибок канала. DNN реализована с помощью алгоритма обучения Левенберга – Маргардта. Экспериментально демонстрируется, что DNN не только обеспечивает более качественные восстановленные изображения, но и менее вычислительная мощность по сравнению с Зональным кодированием DCT, Пороговым кодированием DCT, Set Partitioning in Hierarchical Trees (SPIHT). Предлагается DNN с улучшенной скоростью исправления ошибок канала. Результаты экспериментов указывают на то, что реализованная сеть обеспечивает превосходную способность исправления

ошибок, передавая двоичные изображения через зашумленные каналы с использованием кодирования.

В работе [5] при сжатии изображения с использованием искусственной нейронной сети с прямой связью, обученной с помощью алгоритма обратного распространения вейвлет-преобразования, предлагается сжимать изображения высокого качества. В этом новом подходе прямая связь с тремя скрытыми слоями сеть (FFN) применяется непосредственно как основное сжатие алгоритма компрессии изображений. После обучения с достаточным количеством образцов изображений процесс сжатия будет выполняться на целевом изображении. Веса связи и значения активации каждого нейрона в скрытый слой будут сохранены после обучения. Сжатие достигается за счет использования меньшего количества скрытых нейронов по сравнению с количеством пикселей изображения из-за меньшего количества информации.

В статье [6] предлагается альтернативный подход к сжатию HSI с помощью генеративной нейронной сети (GNN), которая изучает распределение вероятностей реальных данных из случайного скрытого кода. Это достигается путем определения семейства плотностей и нахождения этим семейством и реальным распределением данных. Тогда хорошо обученная нейронная сеть является представлением HSI, а степень сжатия определяется сложностью GNN. Более того, скрытый код можно зашифровать, вставив цифру со случайным распределением, что делает код конфиденциальным. По сравнению с другими алгоритмами, он имеет лучшую производительность при высокой степени сжатия, и есть еще много возможностей для улучшений наряду с быстрым развитием методов глубокого обучения, а степень сжатия определяется сложностью GNN. Экспериментальные примеры представлены, чтобы продемонстрировать потенциал GNN для решения проблем сжатия изображений в области HSI.

Сжатие информации достигается путем обработки более короткого изображения в исходной информации, которое должно содержать аналогичную информацию, но с меньшей длиной иллюстрацией. В статье [7] показано, что нейронные сети являются многообещающим инструментом для сжатия информации, но не для потери данных. Присутствует тенденция смешивать нейронные сети, стандартные методы статистического сжатия, такие как кодирование Хаффмана и арифметическое кодиро-

вание. В работе используются в основном методы, основанные на искусственной нейронной сети, предлагающие новые способы более безопасного сжатия информации в передатчике и декомпрессии в приемнике.

Задача сжатия гиперспектральных данных дистанционного зондирования Земли рассмотрена в работе [8]. Авторы предлагают алгоритм увеличения степени сжатия, что достигается за счет использования формирования вспомогательных данных с высокой избыточностью на основе их байтового представления и учета межканальной корреляции, что, в свою очередь, значительно увеличивает и время обработки снимка.

Для некоторых задач ДЗЗ могут применяться и методы сжатия с потерями. Так, в работе [9] описывается метод на основе многослойного персептрона для компрессии изображений различной конфигурации. Показатели сжатия от количества слоев и количества нейронов в них были показателями при анализе структуры ИНС. Полученный коэффициент сжатия 4.0 достигается при 16 нейронах в скрытом слое ИНС, при этом отмечается, что разработанная нейронная сеть дает меньшие потери при сжатии изображений с низкой резкостью и контрастностью.

Компессию данных на фрагментах любого изображения рассматривают авторы работы [10]. В теории авторы сжимают преобразование с помощью нейросети.

На сегодняшний день не существует программно-инструментального комплекса организации вычислений, обладающего одновременно следующими важнейшими характеристиками:

- кроссплатформенность;
- производительность при ограничениях на количество используемых аппаратных ресурсов.

Результаты исследования и их обсуждение

1. Сжатие данных ДЗЗ с помощью RAQ

Архиватор RAQ применялся в качестве метода сжатия данных, который основан на предсказании частичным совпадением (partial match) и контекстном моделировании (context mixing model). RAQ является универсальным алгоритмом для сжатия любых данных. В наших экспериментах использовались снимки ДЗЗ.

Фрагменты результатов сжатия с помощью RAQ представлены в таблице. Сжатие производилось на персональном компьютере (процессор Intel Core i3-8300).

Фрагменты результатов сжатия данных ДЗЗ

Входные данные ДЗЗ	Время обработки (с)	Результаты сжатия
 <p>ВМР – 34,7 Мб, 4001x3032 пикселей</p>	60,6	ВМР – 16,8 Мб
 <p>ВМР – 6,17 Мб, 1800x1200 пикселей</p>	16,3	ВМР – 2,7 Мб

Окончание таблицы		
Входные данные ДЗЗ	Время обработки (с)	Результаты сжатия
 BMP – 4,11 Мб, 1200x1200 пикселей	11,3	BMP – 1,7 Мб
 BMP – 14,6 Мб, 2493x2048 пикселей	32,4	BMP – 4,2 Мб
 BMP – 11,9 Мб, 2200x1900 пикселей	28,9	BMP – 3,5 Мб
 BMP – 9,9 Мб, 2055x1684 пикселей	25,4	BMP – 2,9 Мб

Эксперименты показали, что RAQ может сжимать без потерь снимки ДЗЗ с коэффициентом сжатия в диапазоне 2.06-3.47 (зависит от исходного изображения).

Заключение

Исследованы алгоритмы сжатия больших потоков целевых данных для увеличения эффективности их передачи, хранения и дальнейшего анализа, показывающие, что в сегодняшнее время не существует программно-инструментального комплекса организации вычислений, обладающего одновременно такими важнейшими характеристиками, как кроссплатформенность и производительность при ограничениях на количество используемых аппаратных ресурсов.

Предложен метод сжатия ДЗЗ без потерь на основе архиватора RAQ. Проведенные эксперименты показали способность данного метода сжимать снимки ДЗЗ с коэффициентом сжатия 3.47. Разработанное алгоритмическое и программное обеспечение можно интегрировать в наземные и бортовые системы летательных аппаратов для повышения эксплуатационных характеристик.

Работа выполнена в рамках Программы фундаментальных исследований Президиума РАН «Фундаментальные основы прорывных технологий в интересах национальной безопасности» (проект «Разработка и исследование методов и технологии высокопроизводительного сжатия целевой информации, передаваемой по каналам космической связи в интересах национальной безопасности Российской Федерации») и проекта РФФИ № 18-29-03011-мк «Исследование и разработка новых методов и технологий для задач интеллектуального анализа и оптимизации обработки больших

потоков данных дистанционного зондирования Земли».

Список литературы

1. Дудин Е.А., Карин С.А., Григорьев А.Н. Сжатие многоспектральных данных дистанционного зондирования Земли с использованием метода главных компонент // Информатика и космос. 2014. № 4. С. 77–81.
2. Аль-Бахдили Х.К., Цветков В.Ю., Конопелько В.К. Сжатие изображений дистанционного зондирования Земли на основе вероятностного кодирования длин серий пикселей // Доклады Белорусского государственного университета информатики и радиоэлектроники. 2017. № 1. С. 65–70.
3. Петров Е.П., Харина Н.Л., Сухих П.Н. Метод сжатия изображений в системах ДЗЗ без потерь // Вестник Самарского государственного аэрокосмического университета имени академика С.П. Королева. 2016. Т. 15. № 2. [Электронный ресурс]. URL: <http://jmla.org/papers/doc/2015/no12/Petrov2015Compression.pdf> (дата обращения: 22.03.2021).
4. Yijing Z., Watkinsa, Mohammad R. Sayeha. Image Data Compression and Noisy Channel Error Correction Using Deep Neural Network – Procedia Computer Science 95 (2016) 145–152. [Electronic resource]. URL: <https://core.ac.uk/download/pdf/82785581.pdf> (date of access: 22.03.2021).
5. Shukla S., Srivastava A. Compression of medical images using feed-forward neural network with LWT. International Journal of Engineering and Technical Research (IJETR). 2018. С. 34–37. [Electronic resource]. URL: <https://media.neliti.com/media/publications/264825-compression-of-medical-images-using-feed-3d0e44ef.pdf> (date of access: 22.03.2021).
6. Deng C., Cen Yi, Zhang L.L. Learning-Based Hyperspectral Imagery Compression through Generative Neural Networks. Remote Sens. 2020. DOI: 10.3390/rs12213657.
7. Kumar Goar V. New Ways for The Compression of Data Using Artificial Neural Network for Transmission – International Refereed Journal of Reviews and Research. 2018. [Electronic resource]. URL: <http://ijrr.com/ijrr/January2018/7.pdf> (date of access: 22.03.2021).
8. Замятин А.В., Сарина А.Ж. Алгоритм сжатия гиперспектральных аэрокосмических изображений с учетом междиапазонной корреляции // Прикладная информатика. 2013. № 5. С. 35–42.
9. Лёзин И.А., Соловьёв А.В. Сжатие изображений с использованием многослойного перцептрона // Известия Самарского научного центра РАН. 2016. № 4–4. С. 770.
10. Сирота А.А., Дрюченко М.А. Обобщённые алгоритмы сжатия изображений на фрагментах произвольной формы и их реализация с использованием искусственных нейронных сетей // Компьютерная оптика. 2015. № 39. С. 751–761.