

---

---

## LOSSLESS-МЕТОД СЖАТИЯ ТОМОГРАФИЧЕСКИХ ДАННЫХ

Сергей Удовенко, Анатолий Шамраев, Елена Шамраева

**Аннотация:** В работе предложен комплексный метод сжатия томографических данных без потерь. Предложенный метод основан на классических и модифицированных алгоритмах сжатия данных без потерь, позволяющих уменьшить не только межэлементную избыточность данных на каждом снимке, но и межкадровую избыточность, обусловленную особенностями томографических данных.

**Ключевые слова:** томограммы, сжатие, модифицированный алгоритм кодирования длин серий, кодирование Хаффмана, межэлементная и межкадровая избыточности.

**ACM Classification Keywords:** I.4 Image Processing And Computer Vision, I.4.2 Compression (Coding), E.4 Coding And Information Theory – Data compaction and compression

---

### Введение

В настоящее время в медицине одним из наиболее информативных диагностических методов является компьютерная томография (рентгеновская компьютерная томография и магнито-резонансная томография). Томография – метод неразрушающего послойного исследования внутренней структуры объекта, основанный на измерении разности ослабления рентгеновского излучения различными по плотности тканями (рентгеновская компьютерная томография) или на измерении электромагнитного отклика ядер атомов водорода на возбуждение их определенной комбинацией электромагнитных волн в постоянном магнитном поле высокой напряженности (магнито-резонансная томография). Применение той или иной разновидности метода диагностики обуславливается характером визуализируемых объектов. Результат исследований представляет собой набор (от полусотни до нескольких сотен) полутоновых изображений (томографических снимков) размером 512×512 пикселей.

На сегодняшний день серьезной проблемой является долгосрочное хранение результатов томографических исследований для оценки эффективности проводимых лечебных процедур и контроля состояния пациента в динамике (полный набор томограмм одного исследования может занимать порядка 100 МБ памяти (размер одного снимка от 500 кБ до 700 кБ)). Решить эту проблему можно путем применения Lossless-методов сжатия данных (без потерь, т.е. со 100% восстановлением исходных данных), т.к. именно этот вариант сжатия является единственно допустимым способом сокращения объема медицинских данных. Существующие универсальные алгоритмы сжатия данных без потерь (например, кодирование Хаффмана, кодирование длин серий (RLE), алгоритмы серии LZ и т.д.) и алгоритмы сжатия изображений без потерь (например, JPEG 2000 в режиме сжатия без потерь, JPEG-LS и др.) [Гонсалес 2006; Salomon 2002; Сергеенко, 2012] не могут быть эффективно использованы для сжатия томографических данных, т.к. не учитывают их специфику, вследствие чего не может быть достигнута высокая степень компрессии.

Специфичность томографических данных заключается в том, что, во-первых, размер всех томографических снимков составляет 512×512 пикселей, а во всех файлах снимки одинаково позиционированы; во-вторых, часть снимка состоит из длинных серий одинаковых (черных) пикселей; в-третьих, различия между двумя соседними томограммами незначительны (рис.1). Если существующие

алгоритмы (например, RLE) учитывают вторую особенность томографических снимков, то первая и третья не принимаются во внимание. Поэтому задача разработки методов компрессии без потерь для сжатия томографических данных является достаточно актуальной.

### Модификация методов сжатия без потерь под томографические данные

Применение алгоритма кодирования длин серий может решить проблему межэлементной избыточности на каждом томографическом снимке путем кодирования серий одинаковых пикселей. Суть алгоритма заключается в преобразовании значений элементов вдоль строки изображения  $f(x, 0), f(x, 1), \dots, f(x, N - 1)$  в набор следующих пар [Гонсалес 2006]:

$$rle = (g_1, w_1), (g_2, w_2), \dots, (g_i, w_i), \dots, \quad (1)$$

где  $g_i$  – значение яркости на отрезке (серии)  $i$ ,  $w_i$  – длина данной серии.

Данное преобразование чаще всего используют для бинаризованных изображений [Salomon 2002].

Однако томографический снимок (рис.1,а) является полутоновым изображением, причем, если черные (с нулевой яркостью) пиксели можно заменить достаточно длинными сериями (вплоть до 512 пикселей), то остальные (ненулевые) пиксели представлены практически единичными значениями длин серий. Поэтому применение метода RLE в виде (1) может быть неэффективно для сжатия томографических данных: все преимущества сокращения кодирования черных пикселей утрачиваются из-за увеличения кода для всех остальных пикселей.

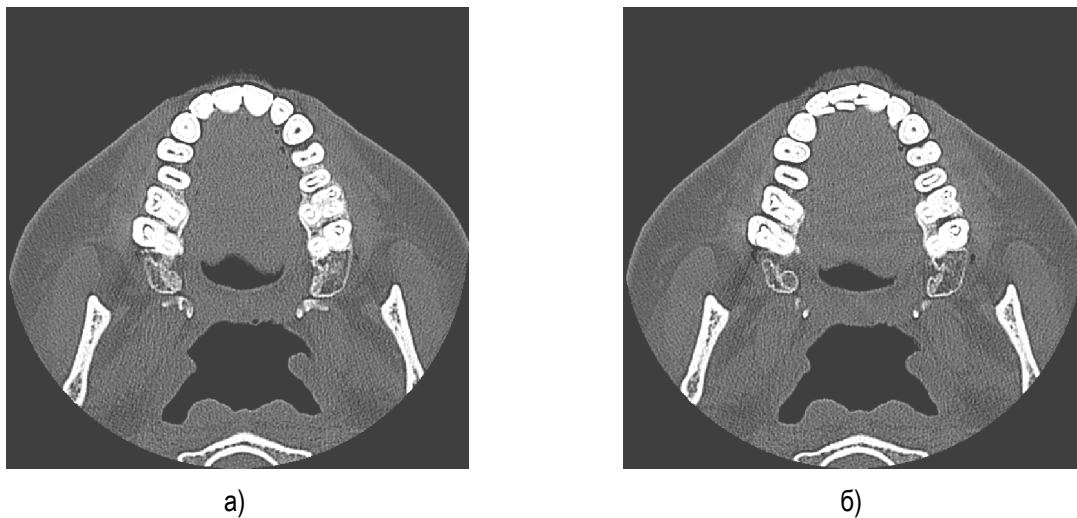


Рис1. Томограммы: а) первый снимок; б) второй снимок

В связи с этим, авторами предлагается модификация метода кодирования длин серий (1), заключающаяся в следующем.

Значения элементов вдоль строки томографического изображения  $f(x, 0), f(x, 1), \dots, f(x, 511)$  преобразуются следующим образом:

$$rle\_m = \begin{cases} (0, w_1), (0, w_2), \dots, (0, w_i), \dots & \text{для пикселей с } g_i = 0; \\ g_1, g_2, \dots, g_i, \dots & \text{в остальных случаях.} \end{cases} \quad (2)$$

---

---

В серии объединяются только пиксели с нулевыми значениями. Меткой начала серии пикселей с нулевой яркостью является нулевое значение элемента ряда, за которым следует значение длины серии  $(0, w_j)$ . Если за значением длины серии следует ненулевое значение, то оно соответствует значению яркости единичного пикселя  $(g_j)$ , за которым может идти как пара значений  $(0, w_j)$ , так и значение  $(g_j)$ . Размер томографических снимков постоянен и равен  $512 \times 512$  пикселей, что подразумевает работу алгоритма сжатия с двухбайтовыми указателями. Для работы с единой структурой данных алгоритм модифицирован под работу с однобайтовыми указателями, для чего на начальном этапе:

- Размер томограммы уменьшается до  $510 \times 510$  путем удаления по ее периметру по одному пикселю (обычно с нулевой яркостью);
- Томограмма по вертикали и горизонтали делится на две равные части и дальнейшее преобразование осуществляется над снимками размером  $255 \times 255$ .

При декомпрессии данных размер томограммы восстанавливается до исходных  $512 \times 512$  пикселей.

С помощью модифицированного RLE-алгоритма (2) устраняется межэлементная избыточность на снимке, однако не устраняется межкадровая избыточность, присутствующая в наборе томографических снимков одного исследования из-за значительной корреляции (95-99%) двух соседних томограмм. Специфичность томографических данных позволяет учесть и устранить межкадровую избыточность с помощью алгоритмов, применимых для компрессии видеоинформации [Ричардсон, 2005], тем самым значительно и без потерь уменьшая объем данных.

Авторами предложен комплексный метод сжатия томографических данных без потерь, состоящий из следующих этапов:

- 1) Базовый (первый) снимок преобразуется с помощью предложенного модифицированного RLE-алгоритма (2);
- 2) Каждый последующий снимок кодируется по следующему принципу кодирования видеоданных: определяется и кодируется расхождение (разность) между текущим и базовым снимками; если новый снимок сильно отличается от базового (рекомендованное значение предельного расхождения – 10%), то его необходимо кодировать независимо с помощью алгоритма (2); новый снимок становится базовым.

Применение разработанного комплексного метода для сжатия томографических данных позволило сократить объем данных в среднем в 4,38 раза, что для методов сжатия без потерь [Устюжанин, 2008] является высоким показателем.

---

## Заключение

В работе предложен комплексный метод компрессии данных без потерь, адаптированный к сжатию томографических данных. Модификация классических алгоритмов позволила удалить как межэлементную избыточность на каждом снимке, так и межкадровую избыточность, присущую томографическим данным. Предложенный метод позволил сократить объем данных в среднем в 4,38 раза.

---

## Литература

[Salomon, 2002] D. Salomon. "Data compression methods", D. Salomon, Springer-Verlag, New York, 2002

[Гонсалес, 2006] Р. Гонсалес. Цифровая обработка изображений. Р. Гонсалес, Р. Вудс. Техносфера, Москва, 2006.

- 
- [Ричардсон, 2005] Я.Ричардсон. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения. Я.Ричардсон. Техносфера, Москва, 2005.
- [Сергеенко, 2012] В.С. Сергеенко. Сжатие данных, речи, звука и изображений в телекоммуникационных системах. В.С. Сергеенко, В.В. Баринев. Радиософт, Москва, 2012.
- [Устюжанин, 2008] Д.В. Устюжанин. Оценка эффективности алгоритмов сжатия МР-изоб/вращений для передачи по компьютерным сетям для телерадиологических консультаций. Д.В. Устюжанин, О.С. Пьяных. Медицинская визуализация. Материалы 2-го Всероссийского национального конгресса по лучевой диагностике и терапии. Москва, 2008. – С.292-293

---

### Информация об авторах

---



**Удовенко Сергей** – д.т.н., профессор Харьковского национального университета радиозлектроники; пр. Ленина 14, 61166, Харьков, Украина; e-mail: [udovenko@kture.kharkov.ua](mailto:udovenko@kture.kharkov.ua).

Основные области научных исследований: интеллектуальные методы обработки информации



**Шамраев Анатолий** – к.т.н., доцент Харьковского национального университета радиозлектроники; пр. Ленина 14, 61166, Харьков, Украина; e-mail: [shamraev@kture.kharkov.ua](mailto:shamraev@kture.kharkov.ua)

Основные области научных исследований: сжатие данных, встраиваемые системы управления



**Шамраева Елена** – к.т.н., доцент Харьковского национального университета радиозлектроники; пр. Ленина 14, 61166, Харьков, Украина; e-mail: [shamraeva.elena@kture.kharkov.ua](mailto:shamraeva.elena@kture.kharkov.ua)

Основные области научных исследований: обработка медицинских изображений, сжатие данных

### Lossless Tomography Data Compression Method

Sergei Udovenko, Anatolii Shamraiev, Olena Shamraieva

**Abstract:** This paper presents a complex compression method of tomographic data without loss. The proposed method is based on the classic and modified lossless data compression algorithms, which allow reducing not only the inter-element data redundancy on each shot, but the inter-frame redundancy, due to the peculiarities of tomographic data.

**Keywords:** tomography, compression, modified algorithm-length coding, Huffman coding, inter-element and inter-frame redundancy.