

МЕТОД СЖАТИЯ ИЗОБРАЖЕНИЙ БЕЗ ПОТЕРЬ С ПОМОЩЬЮ КОНТЕКСТНОГО КОДИРОВАНИЯ ПО ДВОИЧНЫМ УРОВНЯМ

Н. Д. Егоров^а, аспирант

Д. В. Новиков^а, программист, инженер

М. Р. Гильмутдинов^а, канд. техн. наук, доцент

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, РФ

Постановка проблемы: визуальные данные в исходном виде занимают весомый объем дискового пространства, при этом данные изображения ввиду важности содержащейся в них информации требуется хранить с применением таких методов, которые не приведут к появлению на них никаких искажений. Для решения этой задачи используются методы сжатия изображений без потерь. **Цель:** разработка эффективного метода сжатия изображений без потерь. **Результаты:** разработан метод сжатия изображений без потерь с помощью контекстного кодирования по двоичным уровням. Доказано, что лежащий в основе данного метода сжатия способ кодирования изображения по двоичным уровням унарного кода не приводит к увеличению энтропии источника. На известных тестовых множествах изображений проведен сравнительный анализ эффективности сжатия данных разработанным методом и алгоритмами JPEG-LS и JPEG-2000 для сжатия изображений без потерь. Получаемые с помощью предложенного алгоритма файлы в среднем на 16,53 и 6,84 % меньше, чем файлы, получаемые с помощью алгоритмов JPEG-LS и JPEG-2000 соответственно. **Практическая значимость:** разработанный метод позволит сэкономить ресурсы, требуемые для хранения визуальных данных, ввиду более компактного представления этих данных.

Ключевые слова — сжатие изображений, сжатие без потерь, кодирование ошибок предсказания, контекстное моделирование, бинаризация.

Введение

Постоянное улучшение качества (повышение разрешения и числа бит, используемого для представления пикселя) и увеличение количества получаемых визуальных данных (изображений) приводят к росту затрат на их хранение — построению новых центров хранения данных. В связи с этим задача компактного представления визуальных данных особенно актуальна в последние годы. Для ее решения было предложено два основных подхода [1, 2]:

— сжатие без потерь — вся информация, содержащаяся на исходном изображении, представлялась в более компактном виде за счет применения различных методов кодирования источника;

— сжатие с потерями — для повышения эффективности сжатия допускалось частичное уничтожение или искажение информации, содержащейся в исходном изображении.

В данной работе исследуются способы сжатия различных классов изображений, относящихся к общему классу высокоточных данных. Для подобных изображений потеря и искажение хранимых в них данных являются критичными, поэтому методы сжатия с потерями не рассматриваются.

Сжатие без потерь осуществляется за счет анализа статистики кодируемого источника. Задача оценки статистики источника осложнена тем, что большинство источников визуальных данных содержат внутренние зависимости (являются ис-

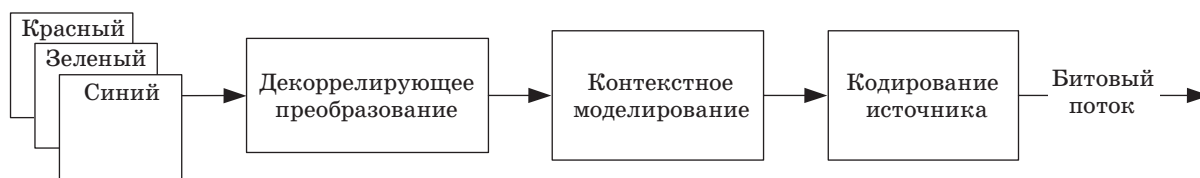
точниками с памятью). Для упрощения решения данной задачи используется следующий подход: сначала к источнику применяется декоррелирующее преобразование для учета памяти источника (устранения имеющейся избыточности), затем — один из разработанных методов кодирования источника без памяти [3, 4]. Типовая схема сжатия изображений на основе данного подхода (рис. 1) включает решение нескольких подзадач:

1) декорреляции данных — устранения присутствующей в данных источника избыточности;

2) контекстного моделирования — разбиения исходного потока данных на подпотоки для оценки статистики источника с учетом оставшихся после предыдущего шага зависимостей;

3) кодирования источника — оценки распределения источника и построения наиболее эффективного кода для компактного представления его (источника) данных. Методы кодирования источника без памяти являются наиболее распространенными и делятся на 2 класса: неадаптивные (предполагают, что входные данные имеют стационарное распределение, не зависящее от времени) и адаптивные.

Для решения первой подзадачи обычно применяются методы, относящиеся к классу алгоритмов Difference Pulse-Code Modulation (DPCM). Текущее обрабатываемое значение предсказывается по набору ранее обработанных данных. Далее вычисляется разность между результатом предсказания \hat{x} и текущим обрабатываемым



■ **Рис. 1.** Типовая схема сжатия изображений
 ■ **Fig. 1.** Typical image compression scheme

значением x : $e = \hat{x} - x$, где e — ошибка предсказания.

В предыдущих работах [5–7] было предложено использовать Active Level Classification Model (ALCM) предсказания [8] в рамках DPCM-класса. Для решения второй и третьей подзадач эффективного кодирования получаемого потока декоррелированных данных (ошибок предсказания) нами использовалась схема, состоящая из двоичного арифметического кодирования и контекстного моделирования по двоичным уровням — Binary Layers Scanning (BLS) [5–7].

В настоящей работе представлены улучшения предложенной схемы, а также доказательство того, что лежащий в основе данного метода сжатия способ кодирования изображения по двоичным уровням унарного кода не приводит к увеличению энтропии источника.

Схема кодирования ошибок предсказания и ее улучшения

Применяемая схема сжатия, впервые введенная в статье [5], предполагает, что цветовые компоненты изображения с размерами W (ширина) на H (высота) уже были обработаны с помощью DPCM-алгоритма. Полученные в результате ошиб-

ки предсказания, также имеющие размерность $W \times H$, обрабатываются независимо.

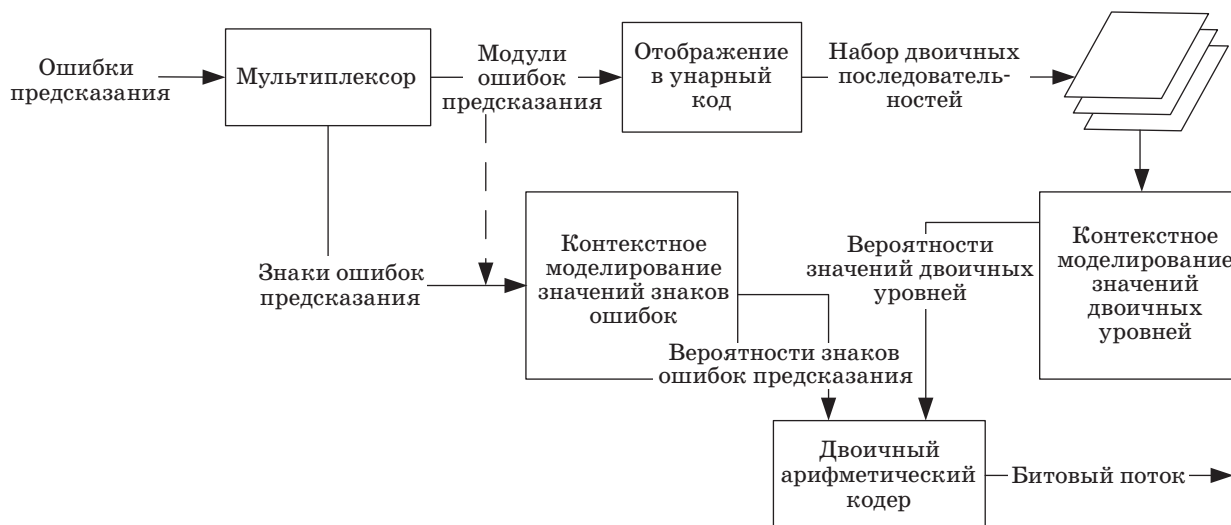
В основе данной схемы лежит идея о применении двоичного арифметического кодера, что помогает добиться универсальности схемы сжатия, доступности ее реализации за счет простоты выполнения оценки вероятности кодируемых символов. Рассмотрим основные особенности схемы кодера (рис. 2).

Раздельная обработка знаков и модулей ошибок предсказания

В качестве входных данных используется множество ошибок выбранного метода предсказания, значения которых лежат в диапазоне $[-I^{\max}, I^{\max}]$, где I^{\max} — максимально возможное значение обрабатываемых визуальных данных (для 8-битного изображения $I^{\max} = 255$). Каждая ошибка предсказания представляется в виде пары последовательно кодируемых значений — модуля ошибки предсказания и знака:

$$\begin{cases} |e_{i,j}| = |e_{i,j}| \\ \text{знак}_{i,j} = \text{знак}(e_{i,j}) \end{cases},$$

где $e_{i,j}$ — ошибка предсказания на позиции i,j ; $|e_{i,j}|$ — модуль ошибки предсказания. Значение



■ **Рис. 2.** Схема кодера BLS
 ■ **Fig. 2.** BLS codec scheme

знака определяется посредством индикаторной функции:

$$\text{знак}(e_{i,j}) = \begin{cases} -1, & \text{если } e_{i,j} < 0 \\ 0, & \text{если } e_{i,j} = 0 \\ 1, & \text{если } e_{i,j} > 0 \end{cases}$$

В предложенной схеме первым кодируется значение модуля ошибки предсказания, что позволяет пропустить кодирование знака в случае, если модуль был равен нулю.

Бинаризация с помощью унарного кода

Применение двоичного арифметического кодера требует бинаризации — процедуры представления символа недвоичного источника в виде последовательности двоичных символов. Для этих целей была выбрана процедура бинаризации с помощью унарного кода:

$$\text{Unary}(x) = (b^{(0)}, b^{(1)}, \dots, b^{(x)}),$$

где x — бинаризуемое значение; $b^{(i)}$ — функция-индикатор:

$$b^{(i)} = \begin{cases} 0, & \text{если } i < x \\ 1, & \text{если } i = x \end{cases}$$

Доказательство того, что метод кодирования изображения по двоичным уровням унарного кода не приводит к увеличению энтропии источника, представлено в следующем разделе.

Способ формирования двоичных уровней для представления модулей ошибок предсказания

Полученный набор $W \times H$ двоичных последовательностей представляют в виде набора дво-

ичных уровней, где k -й двоичный уровень формируется из значений, выбранных с k -й порядковой позиции унарного кода модуля ошибки предсказания. Количество двоичных уровней N будет равно максимальному значению среди всех модулей ошибок предсказания. Значение, находящееся на позиции (i, j) в k -м двоичном уровне, определяется по формуле

$$l_{i,j}^{(k)} = \begin{cases} 0, & \text{если } |e|_{i,j} > k \\ 1, & \text{если } |e|_{i,j} = k \\ -, & \text{если } |e|_{i,j} < k \end{cases}$$

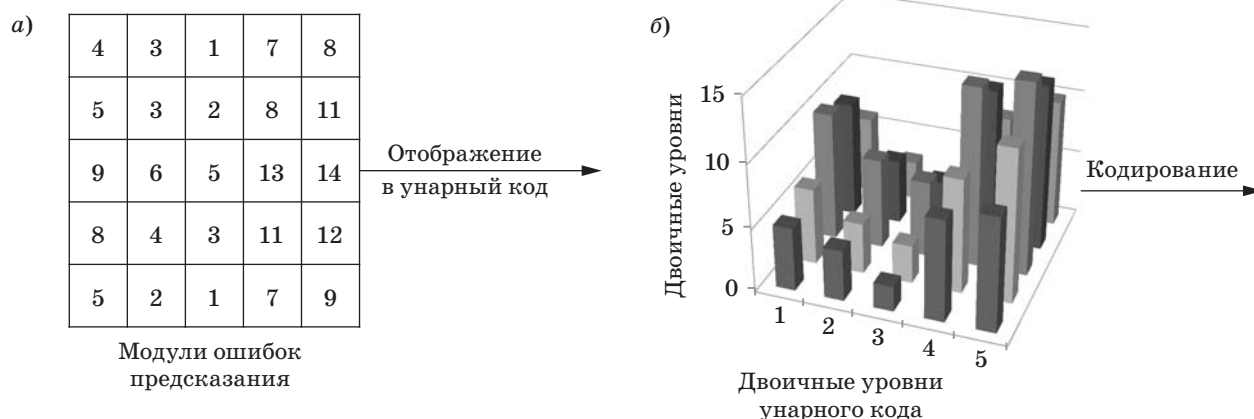
где ‘-’ обозначает то, что в (i, j) в k -м двоичном уровне отсутствует значение и процедура кодирования производиться не будет.

На рис. 3, а представлены значения модулей ошибок предсказания двумерного фрагмента изображения 5×5 ; на рис. 3, б отображен набор двоичных уровней, сформированных из двоичных последовательностей унарного кода.

После формирования двоичных уровней для каждого пикселя производится обработка значений одного двоичного уровня за другим, начиная с 0-го уровня.

Контекстное моделирование знаков и модулей ошибок предсказания

Новый источник данных в виде ошибок предсказания, а соответственно, и получаемые из него двоичные уровни и знаки не являются источниками без памяти. Для учета этого их свойства при кодировании применяется подход, называемый контекстным моделированием. Контекстное моделирование предполагает ведение отдельной статистики (счетчиков частот) в зависимости от



■ **Рис. 3.** Фрагмент изображения 5×5 с соответствующими значениями модулей ошибок предсказания (а) и набор двоичных уровней (б), сформированный с помощью двоичных последовательностей унарного кода
 ■ **Fig. 3.** Image region 5×5 with error module values (а), and set of binary layers (б) formed using binary vectors of unary code

контекста — набора значений, окружающих текущее обрабатываемое значение. Значения, определяющие контекст, называются опорными соседями. Число контекстов напрямую зависит от числа опорных соседей.

С одной стороны, увеличение количества используемых соседей позволяет схеме контекстного моделирования покрывать больше случаев поведения контуров на изображении и повышает эффективность сжатия. С другой стороны, увеличение числа опорных соседей приводит к усложнению схемы кодирования за счет большого количества контекстных моделей. Кроме того, на практике при использовании очень большого количества контекстов в каждый контекст попадает небольшое количество данных, недостаточное для выполнения точной оценки вероятности, приводящее к падению эффективности сжатия.

В работе [6] были определены наиболее релевантные позиции опорных соседей для построения контекстов (рис. 4). Позиции опорных соседей для контекстного моделирования двоичных уровней (рис. 5, а) и знаков (рис. 5, б) были выбраны в соответствии с публикациями различных авторов [9, 10].

Номер контекстной модели Idx , применяемой для обработки значения двоичного уровня с позиции (i, j) , определяется количеством ненулевых значений, присутствующих на позициях используемых опорных соседей, по формуле

$$Idx_{i,j}^{(k)} = \sum_{n,m \in N_{i,j}} I(i+n, j+m)^{(k)},$$

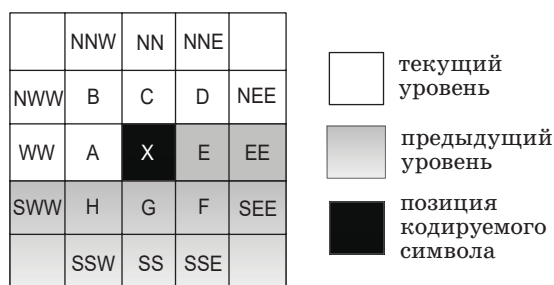
где $N_{i,j}$ — множество позиций используемых опорных соседей для кодирования значения с позиции (i, j) и двоичного уровня k ; $I(.,.)^{(k)}$ — индикаторная функция:

$$I(i, j)^{(k)} = \begin{cases} 1, & \text{если } l_{i,j}^{(k)} = 0 \\ 0, & \text{иначе} \end{cases}.$$

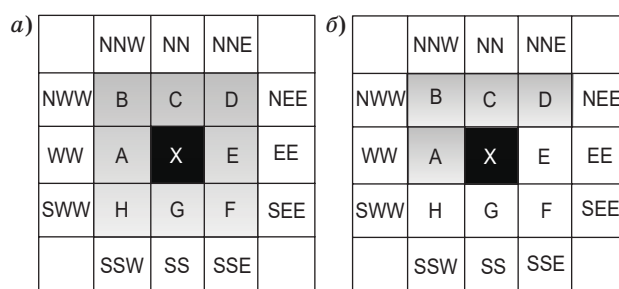
Номер контекстной модели, применяемой для обработки значения знака ошибки с позиции (i, j) , определяется размещением значений знаков с позиций используемых опорных соседей. Для получения уникального номера для каждого возможного размещения значений знаков с используемых позиций применяется формула

$$Idx_{i,j} = \text{sign}((i, j) + \mathbf{a}) + 3 \text{sign}((i, j) + \mathbf{b}) + 9 \text{sign}((i, j) + \mathbf{c}) + 27 \text{sign}((i, j) + \mathbf{d}),$$

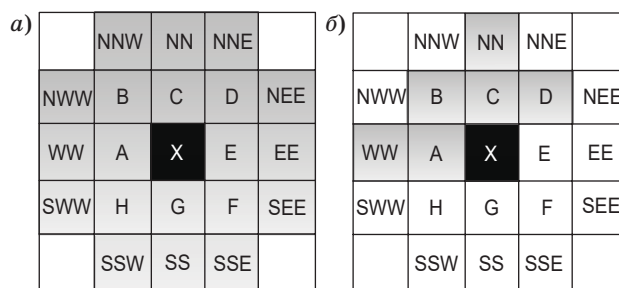
где $\{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}\}$ — векторы, содержащие сдвиги от текущей позиции (i, j) до позиций соседних значений, используемых в контекстном моделировании (см. рис. 5, б).



■ Рис. 4. Набор релевантных позиций соседей при контекстном моделировании
 ■ Fig. 4. Set of neighbor's positions used in context modeling



■ Рис. 5. Набор релевантных позиций соседей при контекстном моделировании двоичных уровней (а) и знаков (б). Серым цветом обозначены позиции используемых соседей
 ■ Fig. 5. Set of neighbor's positions used in context modeling of binary layers (a) and signs (б). Used neighbor's positions are highlighted by grey color



■ Рис. 6. Обновленный набор позиций опорных соседей при контекстном моделировании модулей (а) и знаков (б). Серым цветом обозначены позиции используемых соседей
 ■ Fig. 6. Updated set of neighbor's positions used in context modeling of modules (a) and signs (б). Used neighbor's positions are highlighted by grey color

В данной работе были изменены наборы используемых позиций опорных соседей для построения контекстов двоичных уровней и знаков ошибок предсказания. Новые наборы используемых позиций опорных соседей (рис. 6, а и б) были

подобраны для достижения наибольшей степени сжатия при обработке общедоступных наборов изображений, описанных в последнем разделе. Эффективность обновленной схемы продемонстрирована там же.

В следующем разделе будет доказана эффективность предлагаемого метода сжатия с помощью контекстного кодирования по двоичным уровням.

Применение двоичного арифметического кодирования

Для кодирования источника без памяти на финальной стадии описываемой схемы сжатия используется двоичное арифметическое кодирование. Выбор арифметического кодера [4] объясняется его способностью построения оптимального с точки зрения эффективности сжатия кода для декоррелированных данных при известном распределении. Выбор именно двоичной разновидности арифметического кодирования [11, 12] объясняется практическими соображениями — простотой применения различных контекстных моделей, а также простотой реализации схемы кодирования, адаптивной к изменению статистики кодируемых данных.

Применение двоичного арифметического кодирования требует выполнения оценки распределения кодируемого источника, при этом само распределение источника может быть нестационарным. Для оценки распределения применяется известный подход, предполагающий наличие счетчиков частот обрабатываемых символов. Оценка вероятности появления кодируемого символа производится по формуле Кривичевского — Трофимова [13] для двоичного источника данных, где n_0, n_1 — это счетчики ранее обработанных двоичных символов (0-го и 1-го символа соответственно):

$$p_0 = \frac{n_0 + 1/2}{n_0 + n_1 + 1}; \quad p_1 = \frac{n_1 + 1/2}{n_0 + n_1 + 1}.$$

Стоит отметить, что для устранения остаточной нестационарности в обрабатываемых данных следует учитывать изменение их характеристик с течением времени. Для этого в работе применяется известный подход, называемый экспоненциальным забыванием (Exponential Decaying Machine) [14]. Идея метода заключается в ослаблении влияния ранее обработанных символов и может быть представлена в виде следующей формулы:

$$p_1^{(t+1)} = (1 - \lambda) p_1^{(t)} + \lambda x^{(t)},$$

где λ — коэффициент, управляющий скоростью «забывания»; $p_1^{(t)}$ — оценка вероятности символа

«1» в момент времени t ; $x^{(t)}$ — символ на выходе двоичного источника в момент времени t .

Реализация экспоненциального забывания для обновления счетчиков контекстных моделей представляет собой следующий алгоритм: при обработке порогового количества значений $Reset$ счетчики контекстной модели делятся пополам. Значение счетчика контекстной модели символа x в момент времени t имеет следующую зависимость от ранее обработанных значений:

$$n_{x,t} = \frac{e_{x,1} + \dots + e_{x,Reset}}{2} + e_{x,Reset+1} + \dots + e_{x,2Reset} + \dots + e_{x,t-1},$$

где $e_{x,t}$ — значение, на которое был увеличен счетчик символа x в момент времени t .

Доказательство эффективности предлагаемого метода кодирования по двоичным уровням

В основе предложенного алгоритма сжатия лежит метод бинаризации символов источника X с помощью унарного двоичного кода $Unary(x)$ с последующим кодированием по двоичным уровням $l_{i,j}^{(h)}$. Для доказательства эффективности предложенной схемы покажем, что используемый способ кодирования по двоичным уровням не приводит к возникновению избыточности при кодировании данных, где под избыточностью понимается превышение количества информации, требуемой на представление (хранение) источника, над его энтропией. Наличие данной избыточности обусловлено появлением зависимостей между двоичными потоками, формируемыми с помощью позиций в двоичной последовательности, получаемой после бинаризации.

Согласно известному свойству энтропии [15], обработка информации не приводит к увеличению энтропии:

$$H(Y) \leq H(X),$$

где X — исходный ансамбль; $Y = f(X)$ — ансамбль, полученный из исходного путем применения функции преобразования $f(.)$ ко всем его элементам. В случае если преобразование $f(.)$ является обратимым, данная формула будет являться равенством.

Производимую бинаризацию можно рассматривать как взаимно однозначное отображение исходного недвоичного источника X на множество независимых источников $Y^{(i)}$ ($i = 1, \dots, N$, где

N — количество получаемых двоичных источников). Тогда в соответствии с описанным выше свойством энтропия исходного источника $H(X)$ может быть представлена как взвешенная сумма двоичных источников, полученных после бинаризации:

$$H(X) = \sum_{i=1}^N w_i H(Y^{(i)}),$$

где w_i — вероятность того, что i -й двоичный источник используется при кодировании недвоичного символа источника X . Энтропия исходного недвоичного источника X выражается как

$$H(X) = - \sum_{i=1}^M p_i \log_2(p_i),$$

где M — мощность алфавита источника X ; p_i — вероятность символа i . Энтропия i -го двоичного источника $Y^{(i)}$

$$H(Y^{(i)}) = -p_1^{(i)} \log_2(p_1^{(i)}) - p_2^{(i)} \log_2(p_2^{(i)}),$$

где $p_b^{(i)}$ — вероятность появления символа $b \in \{1, 2\}$ в i -м двоичном потоке.

Несмотря на то что процедура бинаризации является обратимой и энтропия исходного источника $H(X)$ равна взвешенной сумме энтропий, получаемых после бинаризации двоичных источников $w_i H(Y^{(i)})$, использование неэффективной схемы формирования кодируемых источников может приводить к избыточности при их кодировании. В данном разделе будет показано, что формирование кодируемых источников с помощью позиций в двоичной последовательности в случае унарного кода позволяет получить безызбыточное представление, а для равномерного кода такое представление всегда будет избыточным, если распределение источника отлично от равномерного.

Постановка задачи. Докажем, что формирование кодируемых источников с помощью позиций в двоичной последовательности в случае унарного кода позволяет получить безызбыточное представление, т. е. энтропия исходного недвоичного источника X равна взвешенной сумме энтропий получаемых двоичных источников $Y^{(i)}$.

Доказательство: Для выбранного способа бинаризации количество получаемых двоичных источников N равно мощности алфавита исходного недвоичного источника M . Вероятность w_i того, что i -й двоичный источник был использован, будет равна вероятности того, что бинаризуемое число Num , уменьшенное на единицу, больше или равно индексу бинаризуемого источника:

$$w_i = \Pr\{Num - 1 \geq i\} = \sum_{k=i}^M p_k.$$

Тогда распределение i -го двоичного источника $Y^{(i)}$ может быть представлено как

$$\left\{ \begin{aligned} p_1^{(i)} &= \Pr\{Num - 1 = i \mid Num - 1 \geq i\} = \frac{p_i}{\sum_{k=i}^M p_k} \\ p_2^{(i)} &= 1 - p_1^{(i)} = \frac{\sum_{k=i+1}^M p_k}{\sum_{k=i}^M p_k} \end{aligned} \right.$$

Подставим найденные вероятности $p_1^{(i)}$ и $p_2^{(i)}$ в формулу энтропии i -го двоичного источника:

$$H(Y^{(i)}) = - \frac{p_i}{\sum_{k=i}^M p_k} \log_2 \left(\frac{p_i}{\sum_{k=i}^M p_k} \right) - \frac{\sum_{k=i+1}^M p_k}{\sum_{k=i}^M p_k} \log_2 \left(\frac{\sum_{k=i+1}^M p_k}{\sum_{k=i}^M p_k} \right).$$

Взвешенная энтропия i -го двоичного источника

$$\begin{aligned} w_i H(Y^{(i)}) &= \left(\sum_{k=i}^M p_k \right) H(Y^{(i)}) = \\ &= -p_i \log_2 \left(\frac{p_i}{\sum_{k=i}^M p_k} \right) - \sum_{k=i+1}^M p_k \log_2 \left(\frac{\sum_{k=i+1}^M p_k}{\sum_{k=i}^M p_k} \right) = \\ &= -p_i \log_2(p_i) - \sum_{k=i+1}^M p_k \log_2 \left(\sum_{k=i+1}^M p_k \right) + \\ &\quad + \sum_{k=i}^M p_k \log_2 \left(\sum_{k=i}^M p_k \right). \end{aligned}$$

Сумма взвешенных i -го и $(i + 1)$ -го источников равна

$$\begin{aligned} w_i H(Y^{(i)}) + w_{i+1} H(Y^{(i+1)}) &= \\ &= -p_i \log_2(p_i) - p_{i+1} \log_2(p_{i+1}) - \\ &\quad - \sum_{k=i+2}^M p_k \log_2 \left(\sum_{k=i+2}^M p_k \right) + \sum_{k=i}^M p_k \log_2 \left(\sum_{k=i}^M p_k \right). \end{aligned}$$

Из полученного выражения видно, что при сложении взвешенных энтропий i -го и $(i + 1)$ -го источников происходит сокращение второго слагаемого i -го источника и третьего слагаемого $(i + 1)$ -го. Рассмотрим особые случаи в виде взвешенной энтропии первого источника, предпоследнего $(M - 1)$ -го и последнего M -го:

$$w_1 H(Y^{(1)}) = -p_1 \log_2(p_1) - \sum_{k=(2)}^M p_k \log_2 \left(\sum_{k=(2)}^M p_k \right);$$

$$w_{M-1} H(Y^{(M-1)}) = -p_{M-1} \log_2(p_{M-1}) - p_M \log_2(p_M) + \sum_{k=(M-1)}^M p_k \log_2 \left(\sum_{k=(M-1)}^M p_k \right);$$

$$w_M H(Y^{(M)}) = -p_M \log_2(p_M) + p_M \log_2(p_M) = 0.$$

Таким образом, при вычислении суммы взвешенных энтропий не сокращаются только слагаемые вида $p_i \log_2(p_i)$. Тогда итоговая формула суммы взвешенных энтропий $w_i H(Y^{(i)})$ может быть записана как

$$\sum_{i=1}^N w_i H(Y^{(i)}) = -\sum_{i=1}^M p_i \log_2(p_i) = H(X).$$

Как видно из последней формулы, сумма взвешенных энтропий $w_i H(Y^{(i)})$ равна энтропии исходного двоичного источника $H(X)$, что и требовалось доказать.

Постановка задачи. Докажем, что формирование кодируемых источников с помощью позиций в двоичной последовательности в случае равномерного кода приводит к избыточному представлению, если распределение исходного источника X отлично от равномерного.

Разобьем доказательство на две части. В первой части докажем, что в случае если исходный источник X имеет равномерное распределение, формирование кодируемых источников $Y^{(i)}$ с помощью позиций в двоичной последовательности не приведет к увеличению энтропии, т. е. $H(X) = w_i H(Y^{(i)})$. Во второй части докажем, что при распределении исходного источника X , отличным от равномерного, такой способ формирования кодируемых источников будет приводить к избыточному представлению, т. е. $H(X) \neq w_i H(Y^{(i)})$.

Доказательство (часть 1): Считаем, что исходный источник X имеет равномерное распределение:

$$p_i = \frac{1}{M}, \quad i \in [1, M].$$

Также сделаем допущение, что $M = 2^D$, где D — некоторое целое число. Тогда количество

получаемых двоичных источников N равно D , а вес каждого двоичного источника равен единице ($w_i = 1$). Энтропия такого недвоичного источника

$$H(X) = \log_2(M) = D.$$

С учетом сделанного допущения в каждом двоичном источнике $Y^{(i)}$ символ «1» появляется в случае бинаризации одной половины символов исходного источника X , а символ «0» — при бинаризации другой половины. Тогда закон распределения для i -го двоичного источника $H(Y^{(i)})$ может быть представлен в следующем виде:

$$\begin{cases} p_1^{(i)} = \Pr\{p_{Num} \in P^{(i)}(1)\} = \sum_{k \in K(1,i)} p_k \\ p_2^{(i)} = \Pr\{p_{Num} \in P^{(i)}(2)\} = \sum_{k \in K(2,i)} p_k \end{cases},$$

где $p_1^{(i)}, p_2^{(i)}$ — вероятность появления символа «0» и «1» соответственно; $P^{(i)}(1), P^{(i)}(2)$ — множество, содержащее вероятности появления недвоичных символов в исходном источнике X , приводящих к появлению символа «0» и «1» соответственно в i -м двоичном источнике $Y^{(i)}$; $K(b, i)$ — множество индексов вероятностей, содержащихся в множестве $P^{(i)}(b)$.

Зная, что $|K(1, i)| = |K(2, i)| = M/2$ и $p_i = 1/M$ при любом i , упростим выражения для закона распределения i -го двоичного источника $H(Y^{(i)})$:

$$\begin{cases} p_1^{(i)} = \sum_{k \in K(1,i)} \frac{1}{M} = \frac{M}{2} \frac{1}{M} = \frac{1}{2} \\ p_2^{(i)} = \sum_{k \in K(2,i)} \frac{1}{M} = \frac{M}{2} \frac{1}{M} = \frac{1}{2} \end{cases}.$$

Из полученного результата следует, что энтропия любого двоичного источника $H(Y^{(i)})$ равна единице. Тогда сумма взвешенных энтропий двоичных источников $w_i H(Y^{(i)})$ с такими распределениями будет равна

$$\sum_{i=1}^D w_i H(Y^{(i)}) = \sum_{i=1}^D 1 \times 1 = D = H(X).$$

Как видно из полученной формулы, сумма взвешенных энтропий двоичных источников $w_i H(Y^{(i)})$ равна энтропии исходного недвоичного источника $H(X)$ для такого способа формирования двоичных источников, что и требовалось доказать.

Доказательство (часть 2): Для того чтобы распределение исходного источника X отличалось от равномерного, изменим вероятности $(M - 1)$ -го и M -го символов следующим образом:

$$\begin{cases} p_{(M-1)} = \frac{1}{M} + \Delta \\ p_M = \frac{1}{M} - \Delta \end{cases},$$

где Δ — некоторая константа, лежащая в диапазоне $[0; 1/M]$.

То, что мы изменили вероятности конкретных двух символов, никак не скажется на общности доказательства, так как энтропии двух источников, имеющих одинаковые наборы вероятностей символов, равны вне зависимости от порядка следования символов в наборе [15].

С учетом внесенных изменений энтропия исходного недвоичного источника

$$H(X) = - \sum_{i=1}^{M-2} p_i \log_2(p_i) - p_{M-1} \log_2(p_{M-1}) - p_M \log_2(p_M) = \log_2 M - \frac{2}{M} \log_2 M - \left(\frac{1}{M} + \Delta \right) \log_2 \left(\frac{1}{M} + \Delta \right) - \left(\frac{1}{M} - \Delta \right) \log_2 \left(\frac{1}{M} - \Delta \right).$$

Теперь рассмотрим, чему будет равна сумма взвешенных энтропий двоичных источников $w_i H(Y^i)$. Изменения $(M-1)$ -го и M -го символов всегда выражаются в изменении энтропии только одного двоичного источника $Y^{(0)}$. Это объясняется тем, что двоичные коды данных символов при равномерной бинаризации различаются только в одном, младшем, разряде. Для всех остальных двоичных источников $p_{(M-1)}$ и $p_{(M)}$ влияют на появление символа «1». Покажем, что вероятность $p_2^{(i)}$, а соответственно, и $Y^{(i)}$ при $i \neq 0$ не изменится:

$$p_2^{(i)} = \left(\sum_{k \in K(2,i)} \frac{1}{M} \right) + p_{(M-1)} + p_{(M)} = \left(\frac{M-2}{2} \frac{1}{M} \right) + \left(\frac{1}{M} + \Delta \right) + \left(\frac{1}{M} - \Delta \right) = \left(\frac{1}{2} - \frac{2}{M} \right) + \frac{2}{M} = \frac{1}{2}.$$

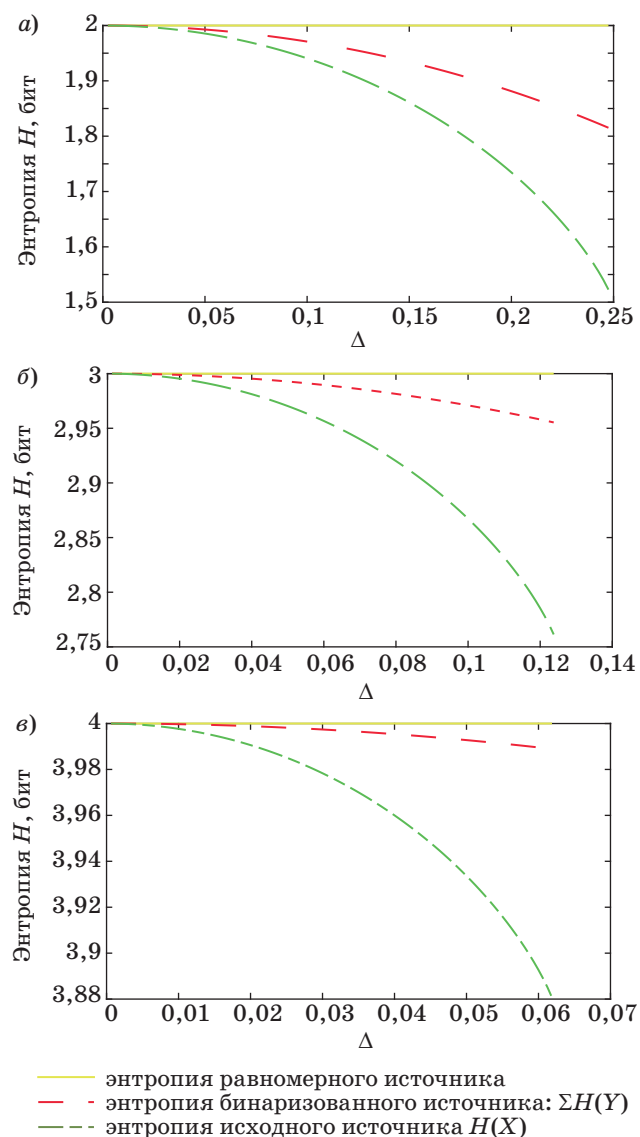
С учетом всех вышеописанных замечаний энтропия взвешенной суммы двоичных источников $w_i H(Y^i)$ может быть представлена в следующем виде:

$$\sum_{i=1}^{\log_2 M} H(Y^i) = \sum_{i=1}^{(\log_2 M)-1} H(Y^i) + H(Y^{(0)}) = \log_2 M - 1 + H(Y^{(0)}) = \log_2 M - 1 - \left(\frac{1}{2} + \Delta \right) \log_2 \left(\frac{1}{2} + \Delta \right) - \left(\frac{1}{2} - \Delta \right) \log_2 \left(\frac{1}{2} - \Delta \right).$$

Дальнейшее доказательство является нетривиальной задачей ввиду наличия множества различных слагаемых в полученных формулах энтропий исходного недвоичного источника $H(X)$ и взвешенной суммы двоичных источников $w_i H(Y^i)$.

Для сравнения данных значений было проведено компьютерное моделирование, согласно которому были зафиксированы значения $M = \{4, 8, 16\}$. Для каждого зафиксированного значения M выбиралось множество из 10^2 значений Δ в диапазоне $[0; 1/M]$. Результаты сравнения энтропии исходного недвоичного источника и взвешенной суммы двоичных источников для различных значений M и Δ показаны на рис. 7, а–в.

Как видно из представленных результатов, сумма взвешенных энтропий двоичных источников $w_i H(Y^i)$ больше энтропии исходного недвоичного источника $H(X)$ в случае его распределения, отличного от равномерного, при данном способе



■ Рис. 7. Энтропии исходного и бинаризованного источников при $M = 4$ (а), $M = 8$ (б) и $M = 16$ (в) и Δ в диапазоне $[0; 1/M]$

■ Fig. 7. Original and binarized sources entropies for $M = 4$ (a), $M = 8$ (b) and $M = 16$ (c) and Δ in range $[0; 1/M]$

формирования двоичных источников, что и требовалось доказать.

Сравнительный анализ эффективности предложенного алгоритма

Проведено сравнение эффективности предложенного алгоритма, его предыдущих версий, а также других общеизвестных алгоритмов [9, 10, 16] на различных общедоступных множествах.

Для выполнения эксперимента были выбраны следующие общедоступные множества 24-битных изображений (табл. 1):

- Lossless Photo Compression Benchmark (LPCB) [17] — содержит множество изображений с высоким разрешением. Изображения относятся к трем различным типам: натуралистичные изображения, орбитальные снимки поверхности Земли, снимки космических объектов и тел;

- Squeeze Chart (SC) [18] Benchmark — содержит небольшое множество изображений различных типов в высоком разрешении;

- Kodak [19] — популярное тестовое множество фотореалистичных изображений;

- Rawzor corpus [20] — содержит фотореалистичные изображения в высоком разрешении;

- Standard — множество из семи таких распространенных изображений, как airplane, baboon, barbara, boats, goldhill, lena, pepper с расширением .bmp.

Для контекстных моделей, используемых при кодировании знаков и модулей ошибок предсказания, на кодируемых тестовых множествах были подобраны наилучшие значения параметра *Reset*: 100 и 500 для знаков и модулей соответственно.

В качестве меры эффективности алгоритмов сжатия рассматривается среднее количество бит, затрачиваемое на компактное представление одного пикселя изображения (24 бита для рассматриваемых цветных изображений). В табл. 2 представлены результаты, полученные при использовании различных схем контекстного моделирования для модулей и знаков ошибок предсказания. Из них можно сделать вывод, что использование

■ **Таблица 1.** Характеристики используемых тестовых множеств

■ **Table 1.** Used test sets characteristics

| Множество | Количество изображений | Размер, Б |
|-----------|------------------------|---------------|
| LPCB | 107 | 3 456 571 880 |
| SC | 7 | 242 932 451 |
| Kodak | 24 | 28 311 912 |
| Rawzor | 14 | 470 611 702 |
| Standard | 7 | 5 982 615 |

■ **Таблица 2.** Результаты сжатия предложенного алгоритма при использовании различных схем контекстного моделирования на различных множествах

■ **Table 2.** Suggested algorithm compression results for different test sets using different context modeling schemes

| Множество | Схема контекстного моделирования | | | |
|-----------|----------------------------------|------------|-------------|----------|
| | Исходная | Для знаков | Для модулей | Итоговая |
| LPCB | 7,913 | 7,892 | 7,89 | 7,869 |
| SC | 10,147 | 10,123 | 10,09 | 10,066 |
| Kodak | 8,611 | 8,596 | 8,569 | 8,554 |
| Rawzor | 9,894 | 9,878 | 9,855 | 9,839 |
| Standard | 12,457 | 12,442 | 12,420 | 12,405 |

■ **Таблица 3.** Результаты сжатия предложенного алгоритма и различных известных алгоритмов на различных множествах

■ **Table 3.** Suggested algorithm and state-of-the-arts algorithms compression results for different test sets

| Множество | Алгоритм сжатия | | |
|-----------|-----------------|---------|-----------|
| | Предложенный | JPEG-LS | JPEG-2000 |
| LPCB | 7,869 | 8,968 | 8,152 |
| SC | 10,066 | 12,18 | 10,52 |
| Kodak | 8,554 | 13,1 | 9,9 |
| Rawzor | 9,839 | 10,64 | 10,62 |
| Standard | 12,405 | 13,91 | 13,12 |

информации от большего количества соседних значений (увеличение количества контекстов) позволяет повысить эффективность предложенного алгоритма, но лишь до определенной степени. При чрезмерном количестве использованных контекстных моделей эффективность алгоритма будет уменьшаться.

В табл. 3 содержатся результаты сжатия без потерь предложенного алгоритма с подобранными наиболее эффективными схемами контекстного моделирования знаков и модулей ошибок предсказания и таких известных алгоритмов сжатия, как JPEG-LS [9, 10] и JPEG-2000 [16]. Из имеющихся результатов можно сделать вывод об эффективности предложенного метода по сравнению с аналогами на различных классах изображений.

Заключение

В работе было доказано, что лежащий в основе предложенного метода сжатия способ кодирования изображений по двоичным уровням унарного кода не приводит к увеличению энтропии. Также было предложено несколько улучшений

для схем контекстного моделирования модулей и знаков ошибок предсказания.

Эффективность самой схемы и предложенных улучшений была продемонстрирована путем сравнения эффективности сжатия с предложенной ранее схемой и другими известными алгоритмами сжатия на известных тестовых множествах изображений. Полученные результаты доказывают эффективность итогового решения и актуальность проделанной работы. В дальнейшем пла-

нируется улучшение предложенного алгоритма путем его настройки для эффективного сжатия специфических классов изображений, например медицинских визуальных данных.

Егоров Н. Д. выражает благодарность за финансовую поддержку Фонду содействия развитию малых форм предприятий в научно-технической сфере, оказанную при выполнении исследования в 2016 г.

Литература

1. **Salomon D.** Data Compression. The Complete Reference. Fourth ed. — London: Springer-Verlag, 2007. — 1092 p. doi:10.1007/978-1-84628-603-2
2. **Ватолин Д., Ратушняк А., Смирнов М., Юкин В.** Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. — М.: ДИАЛОГ-МИФИ, 2002. — 384 с.
3. **Huffman D.** A Method for the Construction of Minimum-Redundancy Codes // Proc. of the IRE. 1952. N 40. P. 1098–1101. doi:10.1109/JRPROC.1952.273898
4. **Witten I. H., Radford N. M., Cleary J. G.** Arithmetic Coding for Data Compression // Communications of the ACM. 1987. Vol. 30. N 6. P. 520–540. doi:10.1145/214762.214771
5. **Gilmutdinov M., Egorov N., Novikov D.** Lossless Image Compression Scheme with Binary Layers Scanning // XIV Intern. Symp. on Problems of Redundancy in Information and Control Systems. 2014. P. 47–51. doi:10.1109/RED.2014.7016703
6. **Egorov N., Novikov D., Gilmutdinov M.** Performance Analysis of Prediction Methods for Lossless Image Compression // Smart Innovation, Systems and Technologies. 2015. Vol. 40. P. 169–178. doi:10.1007/978-3-319-19830-9_16
7. Пат. 2014148851/08 РФ. Способ сжатия изображений (варианты) / М. Р. Гильмутдинов, Н. Д. Егоров, Д. В. Новиков (РФ). — № 2579966; заявл. 04.12.2014; опубл. 10.04.2016, Бюл. № 10. — 2 с.
8. **Speck D.** Fast Robust Adaptation of Predictor Weights from Min/Max Neighboring Pixels for Minimum Conditional Entropy // Signals, Systems and Computers: Conference Record of the Twenty-Ninth Asilomar Conference. 1995. Vol. 1. P. 234–238. doi:10.1109/ACSSC.1995.540547
9. Information Technology — Lossless and Near-lossless Compression of Continuous-Tone Still Images: Baseline // ISO/IEC 14495-1:1999. <https://www.iso.org/standard/22397.html> (дата обращения: 02.04.2017).
10. Information Technology — Lossless and Near-Lossless Compression of Continuous-Tone Still Images: Extensions // ISO/IEC 14495-2:2003. <https://www.iso.org/standard/37700.html> (дата обращения: 02.04.2017).
11. **Marpe D., Schwarz H., Wiegand T.** Context-based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard // IEEE Transactions on Circuits and Systems for Video Technology. July 2003. Vol. 13. N 7. P. 620–636. doi:10.1109/TCSVT.2003.815173
12. **Pennebaker W. B., Mitchell J. L., Langdon G. G., Arps R. B.** An Overview of the basic principles of the Q-Coder Adaptive Binary Arithmetic Coder // IBM Journal of Research and Development. Nov. 1988. Vol. 32. N 6. P. 717–726. doi:10.1147/rd.326.0717
13. **Krichevski R. E., Trofimov V. E.** The Performance of Universal Encoding // IEEE Trans. Inform. Theory. Mar. 1981. Vol. IT-27. P. 199–207. doi:10.1109/TIT.1981.1056331
14. **Meron E.** Universal Finite Memory Prediction, Coding and Estimation of Individual Sequences: Master's thesis. — Tel-Aviv University, 2003. http://www.eng.tau.ac.il/~ronendar/Msc_Thesis.pdf (дата обращения: 16.05.2017).
15. **Кудряшов Б. Д.** Теория информации. — СПб.: СПбГУ ИТМО, 2010. — 188 с.
16. Information Technology — JPEG 2000 Image Coding System: Core Coding System // ISO/IEC 15444-1:2004. <https://www.iso.org/standard/37674.html> (дата обращения: 02.04.2017).
17. Lossless Photo Compression Benchmark by Alexander Ratushnyak. <http://www.imagecompression.info/gralic/LPCB.html> (дата обращения: 02.04.2017).
18. Squeeze Chart Benchmark. <http://www.squeezechart.com/bitmap.html> (дата обращения: 02.04.2017).
19. Kodak Lossless True Color Image Suite. <http://r0k.us/graphics/kodak/> (дата обращения: 02.04.2017).
20. High-Resolution Images Benchmark Presented by Rawzor—Lossless Compression Software for Camera Raw Images. http://www.imagecompression.info/test_images/ (дата обращения: 02.04.2017).

UDC 004.932.2

doi:10.15217/issn1684-8853.2017.6.96

Lossless Image Compression using Binary Layers Scanning Data Encoding

Egorov N. D.^a, Post-Graduate Student, negorov.91@gmail.comNovikov D. V.^a, Programmer, Engineer, dnovikov.suai@gmail.comGilmudtinov M. R.^a, PhD, Tech., Associate Professor, mgilm@vu.spb.ru^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., Saint-Petersburg, 190000, Russian Federation

Introduction: Visual data in raw format take a lot of storage space. Only methods which do not lead to any information loss should be used for the storage of important content. To solve this problem, lossless compression is applied. **Purpose:** Developing an effective algorithm of lossless compression. **Results:** A lossless compression algorithm is developed which uses binary layer scanning data encoding. It has been analytically proved that usage of the proposed encoding scheme with a unary code and binary layers does not increase the entropy of the source. The proposed method was compared with JPEG-LS and JPEG-2000 algorithms on well-known benchmark images. On the average, the files obtained by the new algorithm are 16.53% and 6.84% smaller than those obtained by JPEG-LS and JPEG-2000 respectively. **Practical relevance:** The developed method allows you to save the resources for visual data storage.

Keywords — Image Compression, Lossless Compression, Prediction Error Encoding, Context Modeling, Binarization.

References

- Salomon D. *Data Compression. The Complete Reference*. Fourth ed. London, Springer-Verlag, 2007. 1092 p. doi:10.1007/978-1-84628-603-2
- Vatolin D., Ratushniak A., Smirnov M., Yukin V. *Metody szhatiia dannykh. Ustroistvo arkhivatorov, szhatie izobrazhenii i video* [Image Compression Methods. Universal Archivers, Image and Video Compression]. Moscow, DIA-LOG-MIFI Publ., 2002. 384 p. (In Russian).
- Huffman D. A Method for the Construction of Minimum-Redundancy Codes. *Proc. of the IRE*, 1952, no. 40, pp. 1098–1101. doi:10.1109/JRPROC.1952.273898
- Witten I. H., Radford N. M., Cleary J. G. Arithmetic Coding for Data Compression. *Communications of the ACM*, 1987, vol. 30, no. 6, pp. 520–540. doi:10.1145/214762.214771
- Gilmudtinov M., Egorov N., Novikov D. Lossless Image Compression Scheme with Binary Layers Scanning. *XIV Intern. Symp. on Problems of Redundancy in Information and Control Systems*, 2014, pp. 47–51. doi:10.1109/RED.2014.7016703
- Egorov N., Novikov D., Gilmudtinov M. Performance Analysis of Prediction Methods for Lossless Image Compression. *Smart Innovation, Systems and Technologies*, 2015, vol. 40, pp. 169–178. doi:10.1007/978-3-319-19830-9_16
- Gilmudtinov M., Egorov N., Novikov D. *Sposob szhatiia izobrazhenii (varianty)* [Method of Image Compression (Variants)]. Patent RU, no. 2579966, 2016.
- Speck D. Fast Robust Adaptation of Predictor Weights from Min/Max Neighboring Pixels for Minimum Conditional Entropy. *Signals, Systems and Computers. Conference Record of the Twenty-Ninth Asilomar Conference (1995)*, 1995, vol. 1, pp. 234–238. doi:10.1109/ACSSC.1995.540547
- Information Technology Lossless and Near-Lossless Compression of Continuous-Tone Still Images: Baseline*. ISO/IEC 14495-1:1999. Available at: <https://www.iso.org/standard/37674.html> (accessed 02 April 2017).
- Information Technology Lossless and Near-Lossless Compression of Continuous-tone Still Images: Extensions*. ISO/IEC 14495-2: 2003. Available at: <https://www.iso.org/standard/37674.html> (accessed 02 April 2017).
- Marpe D., Schwarz H., Wiegand T. Context-based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard. *IEEE Transactions on Circuits and Systems for Video Technology*, July 2003, vol. 13, no. 7, pp. 620–636. doi:10.1109/TCSVT.2003.815173
- Pennebaker W. B., Mitchell J. L., Langdon G. G., Arps R. B. An Overview of the basic principles of the Q-Coder Adaptive Binary Arithmetic Coder. *IBM Journal of Research and Development*, November 1988, vol. 32, no. 6, pp. 717–726. doi:10.1147/rd.326.0717
- Krichevski R. E., Trofimov V. E. The Performance of Universal Encoding. *IEEE Trans. Inform. Theory*, March 1981, vol. IT-27, pp. 199–207. doi:10.1109/TIT.1981.1056331
- Meron E. *Universal Finite Memory Prediction, Coding and Estimation of Individual Sequences*. Master's Thesis. Tel-Aviv University, 2003. Available at: <http://www.eng.tau.ac.il/~ronendar/Msc.Thesis.pdf> (accessed 16 May 2017).
- Kudryashov B. *Teoriia informatsii* [Information Theory]. Saint-Petersburg, ITMO Publ., 2010. 188 p. (In Russian)
- Information Technology JPEG 2000 Image Coding System: Core Coding System*. ISO/IEC 15444-1:2004. Available at: <https://www.iso.org/standard/37674.html> (accessed 02 April 2017).
- Lossless Photo Compression Benchmark by Alexander Ratushnyak*. Available at: <http://www.imagecompression.info/gralic/LPCB.html> (accessed 02 April 2017).
- Squeeze Chart Benchmark*. Available at: <http://www.squeezechart.com/bitmap.html> (accessed 02 April 2017).
- Kodak Lossless True Color Image Suite*. Available at: <http://r0k.us/graphics/kodak/> (accessed 02 April 2017).
- High-Resolution Images Benchmark Presented by Rawzor-Lossless Compression Software for Camera Raw Images*. Available at: <http://www.imagecompression.info/test-images/> (accessed 02 April 2017).