

Сравнительное исследование иерархических и дифференциальных методов компрессии изображений

С.А. Денисов¹, М.В. Гашников¹

¹Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

Аннотация. В данной статье проведено сравнительное исследование иерархических и дифференциальных методов компрессии изображений с ограниченной максимальной погрешностью, а именно дифференциальная импульсно-кодированная модуляция (ДИКМ) и иерархическая сеточная интерполяция (HGI). Приведены краткие теоретические сведения об алгоритмах этих методов, результаты работы их реализаций и сравнение этих алгоритмов друг с другом и алгоритмом jpeg.

1. Введение

На сегодняшний день, интернет содержит огромное количество графической информации. Эта информация представлена как в различных видах (чертежи, схемы, рисунки, фотографии), так и в различных форматах (png, jpg, gif, tif, и т.д.) [1]. Такое многообразие графического контента и его количество требуют огромного количества пространства для хранения на серверах. Соответственно, для уменьшения размеров хранимых изображений требуется эффективные алгоритмы компрессии изображений.

Существует довольно большое количество алгоритмов компрессии изображений, но все их можно поделить на 2 больших класса [2, 3]:

1. алгоритмы компрессии с потерями;
2. алгоритмы компрессии без потерь.

Алгоритмы компрессии без потерь имеют более низкий коэффициент сжатия, но при этом позволяют восстановить исходное изображение без погрешностей. Такие алгоритмы применяются там, где качество изображения приоритетнее его размера. Однако, для большинства изображений достаточно применения алгоритмы сжатия с потерями, т.к. при правильном подборе параметров алгоритма получается существенно более высокий коэффициент сжатия при достаточно высоком качестве [4].

Но каким образом подобрать метод компрессии, если важно как качество, так и коэффициент сжатия? Более того, какой алгоритм позволит контролировать вносимую на изображение погрешность? Возможным ответом на эти вопросы могут стать методы компрессии изображений на основе дифференциальной импульсно-кодированной модуляции (ДИКМ) или иерархической сеточной интерполяции (HGI). Данные алгоритмы поддерживают как компрессию без потерь, так и компрессию с потерями, но с заданной максимальной

вносимой погрешности [5, 6]. Данная работа будет посвящена этим методам и их сравнению между собой.

Ниже будут представлены краткие теоретические сведения об исследуемых методах. Затем будет проведена компрессия тестового полутонового однокомпонентного изображения при различных параметрах максимально вносимой погрешности, после чего будет произведено сравнение этих методов между собой и с алгоритмом jpeg.

2. Краткие теоретические сведения о методе компрессии ДИКМ

Дифференциальная импульсно-кодовая модуляция основывается на предположении о наличии корреляции между соседними отсчетами изображения. Значение каждого отсчета x_i оценивается на основе предшествующих x_{i-1}, x_{i-2}, \dots . Оценка \hat{x}_i , полученная в результате, позволяет сформировать разностный сигнал с действительным значением отсчёта $f_i = x_i - \hat{x}_i$. Соответственно, чем выше точность оценки, тем меньшее количество значений принимают отсчёты разностного сигнала, что в свою очередь позволит представлять каждый отсчёт меньшим количеством бит [7].

В общем случае алгоритм состоит из 3 шагов, которые производятся над каждым отсчётом входного изображения:

1. вычисление оценки и формирование разностного сигнала;
2. квантование;
3. восстановление.

После применения алгоритма ко всем отсчётам изображения, производится статистическое кодирования.

Остановимся подробнее на каждом шаге.

2.1. Вычисление оценки и формирование разностного сигнала

Оценка \hat{x}_i , в общем случае, является линейной комбинацией некоторых отсчётов входного изображения $\hat{x}_{i-1}, \dots, \hat{x}_{i-n}$, уже прошедших этап восстановления, и весовых коэффициентов a_1, \dots, a_n , подобранных таким образом, чтобы минимизировать дисперсию разностного сигнала.

$$\hat{x}_{i,j} = \sum_{k=1}^n \sum_{l=1}^m \bar{x}_{i-k,j-l} a_{k,l}$$

В данной работе матрица коэффициентов будет выглядеть следующим образом:

$$\begin{matrix} \mathbf{r} \\ \mathbf{a} \end{matrix} = \begin{pmatrix} -0,5 & 0,5 \\ 1 & 0 \end{pmatrix}$$

Далее вычисляется разностный сигнал.

$$f_{i,j} = x_{i,j} - \hat{x}_{i,j}$$

2.2. Квантование

Далее происходит квантование разностного сигнала при помощи функции квантования $Q(\dots)$.

$$\bar{f}_{i,j} = Q(f_{i,j})$$

При этом неизбежно возникает ошибка квантования.

$$\varepsilon_{i,j} = f_{i,j} - \bar{f}_{i,j}$$

2.3. Восстановление

Для того чтобы данная ошибка не накапливалась в выходном изображении, отсчёт изображения восстанавливается из квантованной ошибки и, в дальнейшем, используется для формирования оценки последующих отсчётов.

$$\bar{x}_i = \bar{f}_i + \sum_{k=1}^n \sum_{l=1}^m \bar{x}_{i-k, j-l} a_{k,l}$$

Как было упомянуто ранее, после применения алгоритма ко всем отсчётам изображения, к изображению, представленному разностным сигналом, применяется статистический кодер. Этим шагом завершается процесс компрессии изображения.

Процесс декомпрессии заключается в применении статистического декодера и процессе восстановления, который эквивалентен тому же процессу при кодировании.

3. Краткие теоретические сведения о методе компрессии HGI

Рассмотрим метод HGI, изначально предназначенный для компрессии однокомпонентного двумерного изображения. Метод HGI основан на специальном иерархическом представлении исходного изображения $X = \{x(m, n)\}$ в виде объединения иерархических уровней X_l :

$$X = \bigcup_{l=0}^{L-1} X_l, (1)$$

$$X_{L-1} = \{x_L(m, n)\}, (2)$$

$$X_l = \{x_l(m, n)\} \setminus \{x_{l+1}(m, n)\}, l < L-1, (3)$$

где L – количество иерархических уровней, а $\{x_l(m, n)\}$ – массив отсчётов изображения, взятых с шагом 2^l по каждой координате.

Представление (1)–(3) даёт возможность при компрессии и декомпрессии обрабатывать уровни последовательно, начиная со старшего уровня X_{L-1} , причём отсчёты менее детального уровня используются для интерполяции отсчётов более детального уровня. Количество уровней выбирается таким образом, чтобы объём старшего (наиболее прореженного) уровня был пренебрежимо мал по сравнению с объёмом данных всего изображения. Это позволяет не сжимать старший уровень [8]. Процедура компрессии любого из следующих уровней $X_l, l < L-1$ состоит из следующих этапов:

1. интерполяция и вычисление постинтерполяционных остатков;
2. квантование постинтерполяционных остатков;
3. восстановление значений отсчетов.

После того как эти этапы завершены, применяется статистическое кодирование.

3.1. Интерполяция

Производится интерполяция отсчётов $\{x_l(m, n)\}$ уровня X_l на основе отсчётов $\{\bar{x}_k(m, n)\}, k > l$ уровней $\{X_k\}, k > l$, уже прошедших квантование и восстановление:

$$\hat{x}_l(m, n) = I\left(\bigcup_{k=l+1}^{L-1} \{\bar{x}_k(m, n)\}\right),$$

где $\{\hat{x}_k(m, n)\}$ – интерполирующие значения отсчётов, а $I(\dots)$ – функция, задающая в общем виде некоторый интерполятор. В данной работе будет использован интерполятор, вычисляющий среднее между ближайшими отсчётами старших уровней.

Вычисляется разность между реальным и интерполирующим значением отсчётов:

$$f_l(m, n) = x_l(m, n) - \hat{x}_l(m, n)$$

3.2. Квантование

Полученные остатки квантуются:

$$\bar{f}_l(m, n) = Q(f_l(m, n))$$

где $Q(..)$ - функция квантования.

3.3. Восстановление значений отсчетов

Производится восстановление отсчетов для дальнейшей работы алгоритма.

$$\bar{x}_l(m, n) = \bar{f}_l(m, n) + \hat{x}_l(m, n)$$

4. Детали реализации алгоритмов

Для больших изображений или изображений с большим разбросом отсчетов яркости имеет смысл производить деление изображение на части. Для каждой части отдельно применяется метод HGI. Таким образом, алгоритм имеет возможность подстроиться под локальные изменения яркости, что может увеличить коэффициент сжатия [9].

В данной работе в качестве функции квантования для алгоритмов будет использоваться квантователь с фиксированной максимальной погрешностью.

$$\bar{f} = \text{sign}(f) \cdot (2\varepsilon_{\max} + 1) \left\lfloor \frac{|f| + \varepsilon_{\max}}{2\varepsilon_{\max} + 1} \right\rfloor,$$

где ε_{\max} - максимальная вносимая погрешность, $\lfloor \dots \rfloor$ - целая часть от деления.

Статистический кодер будет представлен кодером Хаффмана.

5. Результаты работы алгоритмов

На рисунке 1 представлено тестовое изображение, сжатое методом ДИКМ. Даже при довольно большой максимальной погрешности, общее качество изображения (различимость образов, контуров) снижается незначительно, хотя появляется заметный шум.

На рисунке 2 представлено тестовое изображение, сжатое методом HGI. При сравнительно низкой максимальной погрешности (от 0 до 12), HGI вносит в конечное изображение меньше шумов, чем ДИКМ, но при дальнейшем повышении погрешности начинают проявляться заметные артефакты.

6. Сравнение алгоритмов

На рисунке 3 представлен график зависимости коэффициента сжатия от максимально вносимой погрешности. Точки на графиках соответствуют значениям максимальной погрешности от 2 до 20 с шагом 2. Легко увидеть, что HGI имеет немного больший коэффициент сжатия при одинаковой максимальной погрешности.

На рисунке 4 представлен график зависимости коэффициента сжатия от среднеквадратичной ошибки. И здесь HGI обходит ДИКМ, но начиная с СКО = 12, jpeg начинает лидировать по этому показателю. Таким образом, при использовании HGI с максимальной погрешностью меньше или равной 12 можно добиться наилучших показателей среди рассматриваемых алгоритмов.

7. Заключение

В данной работе были исследованы 2 метода компрессии изображения: ДИКМ и HGI. Были представлены теоретические сведения о работе этих методов, а также проведено сравнение этих методов между собой и с алгоритмом jpeg.

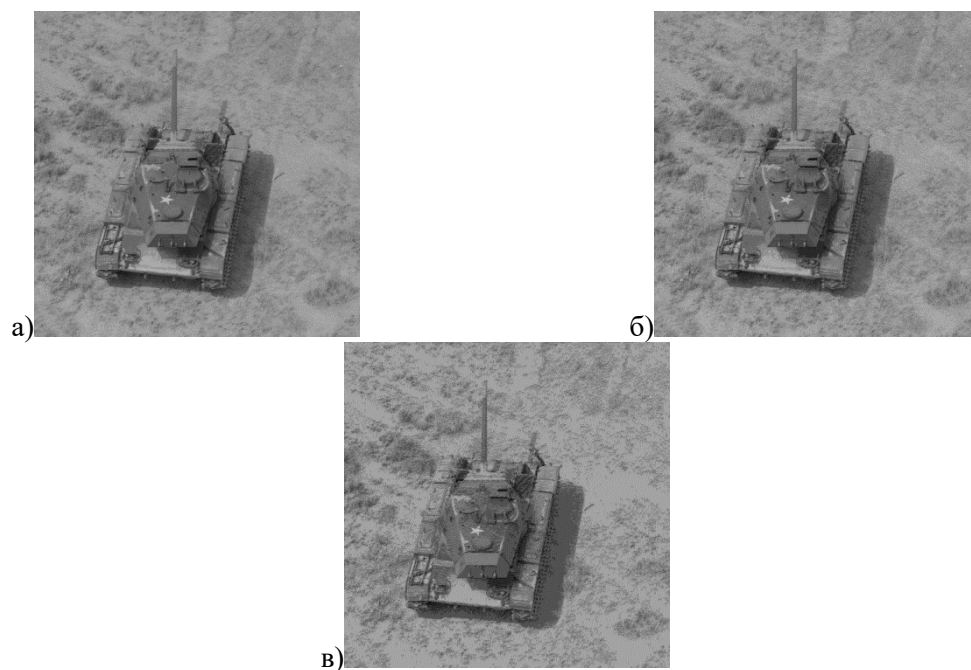


Рисунок 1. Восстановленное изображение после сжатия методом ДИКМ. Максимальная погрешность: а - 2, б - 10, в – 20.

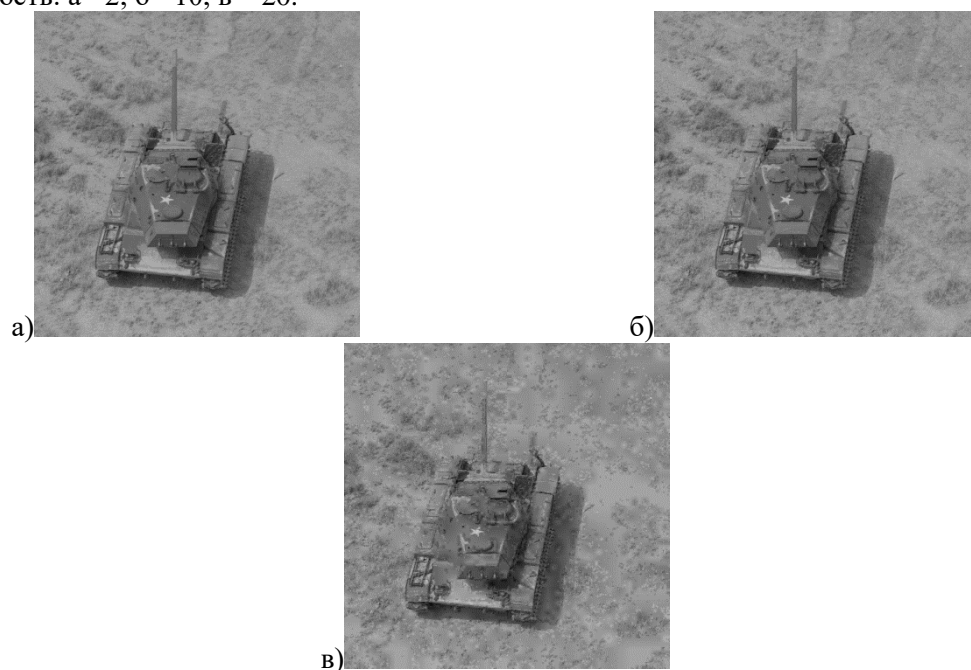


Рисунок 2. Восстановленное изображение после сжатия методом HGI. Максимальная погрешность: а - 2, б - 10, в – 20.

По итогам сравнения данных алгоритмов были сделаны следующие выводы:

1. HGI имеет больший коэффициент сжатия, нежели ДИКМ;
2. HGI меньше зашумляет изображение при небольшой максимальной погрешности, но при этом при её повышении может оставлять заметные артефакты на изображении. ДИКМ, напротив, слабо влияет на качество образа, но вносит больший шум;

3. Оба алгоритма значительно выигрывают jpeg по соотношению коэффициент сжатия/максимальная погрешность, но проигрывают ему по соотношению коэффициент сжатия/СКО при повышении максимальной погрешности;
4. Если максимальная вносимая погрешность не превышает 12, то наилучшим решением будет использование HGI. В противном случае придётся выбирать между jpeg (коэффициент сжатия/СКО) и ДИКМ (коэффициент сжатия/максимальная погрешность).

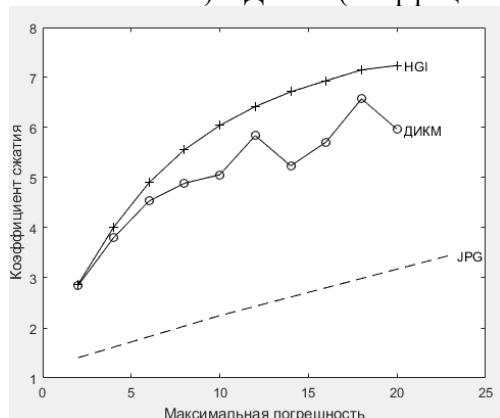


Рисунок 3. График зависимости коэффициента сжатия от максимально вносимой погрешности.

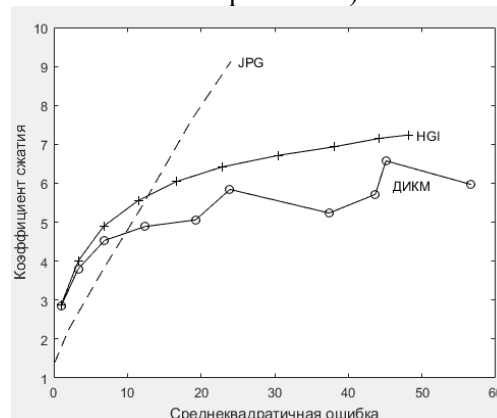


Рисунок 4. График зависимости коэффициента сжатия от среднеквадратичной ошибки.

8. Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 18-01-00667, 18-07-01312.

9. Литература

- [1] Миано, Дж. Форматы и алгоритмы сжатия изображений в действии / Д. Миано. – Москва: Триумф, 2003.
- [2] Смит, С. Цифровая обработка сигналов. Практическое руководство для инженеров и научных работников. / С. Смит. Пер. с англ. А.Ю. Линовича, С.В. Витязева, И.С. Гусинского. – Москва: Додека-XXI, 2011.
- [3] Артюшенко, В.М. Цифровое сжатие видеоинформации и звука: учебное пособие / В.М. Артюшенко, О.И. Шелухин, М.Ю. Афонин; под ред. В. М. Артюшенко. – Москва: Дашков и К, 2003.
- [4] Тропченко, А.Ю. Методы сжатия изображений, аудио-сигналов и видео: учебное пособие / А.Ю. Тропченко, А.А. Тропченко. – Санкт-Петербург: СПбГУ ИТМО, 2009.
- [5] Salomon, D. Data Compression. The Complete Reference. / D. Salomon. – Springer-Verlag, 2007.
- [6] Chang, C. Hyperspectral data exploitation: theory and applications / C. Chang. – Wiley-Interscience, 2007.
- [7] Ватолин, Д. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. – Москва: ДИАЛОГ-МИФИ, 2002.
- [8] Gashnikov, M.V. Compression Method for Real-Time Systems of Remote Sensing / M.V. Gashnikov, N.I. Glumov, V.V. Sergeyev // Proceedings of 15th International Conference on Pattern Recognition. Barcelona, 2000. – Vol.3. – P. 232-235.
- [9] Гашников, М.В. Минимизация энтропии постинтерполяционных остатков при компрессии изображений на основе иерархической сеточной интерполяции / М.В. Гашников // Компьютерная оптика. – 2017. – Т. 41, № 2. – С. 266-275. DOI: 10.18287/2412-6179-2017-41-2-266-275.

Comparative study of hierarchical and differential methods of image

S. Denisov¹, M. Gashnikov¹

¹Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. In this paper, a comparative study of hierarchical and differential methods of compression of images with a limited maximum error, namely, differential pulse code modulation (DPCM) and hierarchical grid interpolation (HGI), is carried out. Brief theoretical information about the algorithms of these methods, the results of the work of their implementations and comparison of these algorithms with each other and the algorithm jpeg are given.

Keywords: hierarchical, image compression, algorithms, DPCM, HGI, differential pulse code modulation, hierarchical grid interpolation, comparative study.