

УДК 004.932.2

ОПТИМИЗАЦИЯ РАЗБИЕНИЯ ИЗОБРАЖЕНИЯ В ФОРМЕ КВАДРОДЕРЕВА ПО КРИТЕРИЮ МИНИМАЛЬНОЙ ДЛИНЫ ОПИСАНИЯ ВО ФРАКТАЛЬНОМ СЖАТИИ

В.В. Окунев, А.С. Потапов

Проведено исследование эффективности разбиения изображения в форме квадродерева применительно к фрактальным методам компрессии. На основе принципа минимальной длины описания введен информационный критерий разделения рангового блока на подблоки. Показано, что модифицированный метод разбиения по квадродереву позволяет достигать одновременного увеличения степени компрессии и качества восстановленного изображения по сравнению с традиционными методами, основанными на критерии порогового значения среднеквадратичного отклонения блоков.

Ключевые слова: фрактальное сжатие, компрессия изображений, минимальная длина описания, квадродерево, оптимизация.

Введение

В настоящее время благодаря значительному распространению электронно-вычислительной техники хранение и обработка данных различного типа осуществляется преимущественно в цифровом виде. Все большую значимость приобретают мультимедийные типы данных – видеофайлы, аудиозаписи, цифровые изображения. В связи с этим одной из актуальных проблем современных информационных технологий представляется разработка эффективных методов компрессии (сжатия) мультимедийных данных, а в особенности – графической информации. Актуальность подчеркивается тем, что технологии обработки, хранения, передачи и защиты информации входят в Перечень критических технологий Российской Федерации.

К мультимедийной информации чаще всего применяется сжатие с потерями. Справедливость такого применения обусловлена тем фактом, что для мультимедийных объектов, как правило, можно отказаться от хранения каких-либо их особенностей (например, мелких деталей на изображении, не воспри-

нимаемых человеческим ухом звуковых частот в аудиозаписи) в пользу увеличения степени компрессии. В действительности существуют распространенные алгоритмы компрессии без потерь и для мультимедийных объектов, такие как FLAC для звуковых файлов или PNG для цифровых изображений. Однако такие форматы проектировались как универсальные для своего типа данных и, как следствие, оказались неспособными учитывать особенности конкретного сжимаемого файла, что в результате привело к существенному проигрышу в степени сжатия по сравнению с аналогичными алгоритмами компрессии с потерями.

Для цифровых изображений как класса мультимедийных данных в настоящее время единственным действительно широко используемым форматом сжатия с потерями является JPEG, получивший широкое применение благодаря распространению цифровых фотоаппаратов, сканеров и других устройств захвата изображений. С учетом количества представленных в формате JPEG изображений очевидны весьма значительные потери, связанные с хранением, передачей и обработкой, возможно, неоптимально (по качеству и степени компрессии) сжатой информации.

С учетом вышесказанного представляется актуальным исследование методов сжатия, основанных на иных представлениях изображений (по сравнению с наиболее распространенными пространственно-спектральными), к которым относятся фрактальные представления, опирающиеся на общее свойство самоподобия изображений. Однако существующие методы фрактального сжатия изображений требуют существенного развития, чтобы их можно было рассматривать в качестве реальной альтернативы JPEG для многих классов изображений, используемых как в научно-технической сфере, так и в повседневной жизни. В частности, остается открытым вопрос о выборе разбиения изображения на ранговые блоки, которое бы обеспечивало максимальное качество восстановленного изображения. В настоящей работе рассмотрена задача оптимизации такого выбора для фрактальных методов компрессии.

Необходимо отметить, что для сохранения практической приемлемости метода фрактального сжатия число сравниваемых разбиений должно расти медленнее, чем линейная функция от размера изображения. Одной из распространенных реализаций такой идеи является разбиение в форме квадродерева [1, 2]. Основная идея такого разбиения – это разделение рангового блока на равные подблоки и повторный поиск оптимальных доменов в случае, если найденное сходство не удовлетворяет какому-либо критерию. Как правило, в качестве такого критерия используется пороговое значение среднеквадратичного отклонения (СКО) рангового блока от оптимального домена. Такой подход нельзя признать универсальным, так как, во-первых, не существует однозначного алгоритма определения данного порогового значения, а во-вторых, учитывается только качество восстановленного изображения, что может привести к ухудшению результирующей степени компрессии. В работе предложен новый критерий разделения на подблоки, лишенный вышеописанных недостатков. Данный критерий базируется на требованиях принципа минимальной длины описания (МДО).

Принцип минимальной длины описания

Основная идея принципа МДО является воплощением широко известного философского правила бритвы Оккама, которое гласит: «То, что можно объяснить посредством меньшего, не следует выражать посредством большего», или «Без необходимости не следует утверждать многое».

Сформулируем принцип МДО следующим образом: «Среди множества моделей следует выбрать ту, которая позволяет минимизировать сумму: 1) длины описания модели (в битах); 2) длины данных, описанных посредством этой модели (в битах)» [3]. Такую формулировку можно считать обобщением нескольких концепций, предложенных разными исследователями в области машинного обучения [4]. Эти концепции отличаются в деталях, однако все они связываются авторами с теоретико-информационной формализацией правила бритвы Оккама.

Приведем некоторые пояснения, которые позволят отождествить понятия, использующиеся в формулировке принципа МДО и в теории фрактального сжатия, или – обобщая – в теории сжатия изображений с потерями. Под *моделью* будем понимать некоторую информационную структуру, которая представляет собой закодированное (сжатое) изображение, полученное в результате работы той или иной разновидности алгоритма компрессии, который в данном случае будет являться представлением изображений. Модель в компактной форме отражает закономерности в исходных данных и позволяет их восстановить с большой точностью. В качестве *длины описания модели* будем рассматривать объем L_{img} сжатого изображения в битах. *Длину данных, описанных посредством модели*, обозначим через L_{loss} . При построении полного описания изображения эту часть данных также необходимо учитывать для получения корректной оценки критерия МДО. В случае сжатия с потерями данный компонент длины описания – это «объем потерь», т.е. количество информации, содержащееся в отклонении восстановленного после сжатия изображения от исходного изображения. Таким образом, принцип МДО обобщает один из стандартных критериев оценивания качества восстановленного изображения и в то же время позволяет построить интегральный критерий эффективности сжатия:

$$L = L_{loss} + L_{img}.$$

Полученный критерий может применяться при оценке качества компрессии изображений с потерями [5]. Для расчета суммарной длины описания сжатого изображения необходимо оценить каждый из компонентов данного критерия.

Для оценки L_{img} рассмотрим, какую информацию необходимо сохранить в сжатом файле, чтобы иметь возможность осуществить восстановление изображения. Если рассматривается один из N_R вариантов разбиения изображения на ранговые блоки, то в среднем будет необходимо $\log_2 N_R$ бит для указания, какое именно разбиение было выбрано. Предположим, что для каждого такого разбиения детерминированным образом устанавливается множество доменных блоков, т.е. отдельно описывать множество доменов не требуется. Далее необходимо указать, какому доменному блоку соответствует каждый ранговый блок. Если число доменных блоков для выбранного варианта разбиения равно n_d , то для кодирования номера доменного блока потребуется $\log_2 n_d$ бит. Также необходимо закодировать коэффициенты преобразования. Пусть число коэффициентов преобразования равно n_p , а на каждый из них нужно выделить в среднем B_{n_p} бит. Таким образом, размер сжатого изображения можно оценить как

$$L_{img} = \log_2 N_R + n_r (\log_2 n_d + n_p B_{n_p}). \quad (1)$$

Оценка L_{loss} может быть произведена следующим образом. Представим разность исходного и восстановленного изображений тоже как изображение, которое может быть каким-либо образом закодировано. Для выполнения простейшей оценки будем полагать, что отклонения в каждом пикселе являются независимыми отсчетами некоторой случайной величины. Если считать, что это предположение нарушается, то в указанной разности должны сохраняться какие-либо регулярные компоненты, которые следовало бы вынести в модель изображения.

Пусть $f(x, y)$ – оригинальное изображение, а $\hat{f}(x, y)$ – восстановленное, тогда для простейшей оценки количества информационных потерь в результате сжатия можно использовать энтропию разностей:

$$L_{loss} = S \cdot H[\hat{f}(x, y) - f(x, y)], \quad (2)$$

где S – площадь изображения; $H(x) = -\sum_{i=0}^{n-1} p_i \log_2 p_i$, $p_i = \frac{h_i}{S}$, h_i – количество пикселей яркости i , а n – количество градаций серого. Таким образом, интегральный критерий эффективности сжатия, построенный на основе принципа МДО, будет иметь следующий вид:

$$L = L_{loss} + L_{img} = S \cdot H[\hat{f}(x, y) - f(x, y)] + \log_2 N_R + n_r (\log_2 n_d + n_p B_{n_p}). \quad (3)$$

Для сравнения качества сжатия изображения при различных параметрах компрессии необходимо для каждого из наборов параметров вычислить длину описания L закодированного изображения. Набор параметров, при котором будет достигаться минимальное значение L , будет обеспечивать оптимальную (одновременно по степени сжатия и качеству восстановленного изображения) компрессию.

Квадродерево

Вышеописанная схема может быть применена для оценки длины описания закодированного изображения в процессе построения квадродерева. Рассмотрим случай, когда изображение разбито на ранговые блоки равных размеров и для каждого рангового блока найден оптимальный домен. После этого по каждому ранговому блоку необходимо проверить, может ли привести его разделение на подблоки к уменьшению длины описания. Такая проверка производится с использованием критерия МДО, при этом слагаемое L_{loss} вычисляется через значение СКО. Вычисление L_{loss} напрямую через энтропию невязок блоков не может быть использовано ввиду того, что применительно к блокам изображения подобный способ дает весьма неточные результаты из-за относительно малого количества пикселей, принадлежащих блоку. Осуществлять декомпрессию изображения на каждом шаге разделения блоков нецелесообразно, поэтому, в частности, не представляется возможным напрямую сравнивать значения СКО оригинального изображения от восстановленного; вместо этого используется оценочное значение СКО, которое вычисляется как среднее арифметическое всех значений СКО ранговых блоков от соответствующих доменов.

Уточним компоненты формулы (3) для случая квадродерева. Пусть все ранговые блоки имеют размер $r \times r$ пикселей, а домены – $d \times d$ пикселей, при этом $d = 2r$. Через r_i будем обозначать i -й ранговый блок. Оценка $\Delta L_{loss}^{(r_i)}$ разности между СКО, соответствующим i -му ранговому блоку, и СКО, соответствующим всему изображению, будет выглядеть так:

$$\Delta L_{loss}^{(r_i)} = 0,75r^2 (\log_2 \overline{\text{RMSE}}(r_i, d_i) - \log_2 \overline{\text{RMSE}}_f), \quad (4)$$

где RMSE – средноквадратичное отклонение (root mean square error), а $\overline{\text{RMSE}}_f$ – оценка СКО восстановленного изображения от оригинального, вычисляемая по следующей формуле:

$$\overline{\text{RMSE}}_f = \frac{1}{n_r} \sum_{j=1}^{n_r} \text{RMSE}(r_j, d_j),$$

где n_r – общее количество ранговых блоков. Коэффициент 0,75 в формуле (4) вводится для обеспечения приближенного равенства оценке (2).

Оценку L_{img} также уточним. Вместо кодирования номера разбиения (из N_R возможных) в случае квадродерева следует хранить для каждого рангового блока информацию о том, разделен ли он на подблоки или нет. Для того чтобы найти величину $\Delta L_{img}^{(r_i)}$, которая показывает изменение длины описания в случае разделения i -го рангового блока на подблоки, необходимо сравнить длину описания сжатого файла до и после разделения, считая подблоки самостоятельными блоками. Так как рассматривается всего один способ разбиения на ранговые блоки, то $\log_2 N_R = \log_2 1 = 0$. В силу того, что в сжатом файле добавляется описание четырех подблоков и удаляется описание «родительского» рангового блока, получим $n_p = 4 - 1 = 3$. В соответствии с формулой (1) для L_{img} , а также с учетом дополнительного бита для указания факта разделения блока, имеем:

$$\Delta L_{img}^{(r_i)} = 3(\log_2 n_d + n_p B_{n_p} + 1). \quad (5)$$

Принимая во внимание формулы (4) и (5), получаем следующий итоговый критерий МДО, позволяющий определить для каждого рангового блока рациональность его разбиения на подблоки:

$$\Delta L^{(r_i)} = \Delta L_{loss}^{(r_i)} + \Delta L_{img}^{(r_i)} = 0,75r^2(\log_2 \text{RMSE}(r_i, d_i) - \log_2 \overline{\text{RMSE}}_f) + 3(\log_2 n_d + n_p B_{n_p} + 1).$$

Справедливость $\Delta L^{(r_i)} \leq 0$ для рассматриваемого рангового блока означает, что его разделение на подблоки увеличит качество сжатия, а именно, улучшит качество восстановленного изображения, а длина сжатого файла возрастет незначительно. Если же имеет место $\Delta L^{(r_i)} > 0$, то следует разделение текущего рангового блока считать нерациональным и переходить к рассмотрению следующего блока.

Предположим, что после рассмотрения конкретного рангового блока было принято решение о разделении его на подблоки. Тогда за СКО, соответствующее этому ранговому блоку, принимается среднее арифметическое от СКО, соответствующее его подблокам, и после этого пересчитывается значение СКО для всего изображения. Подсчеты для дальнейших рассматриваемых ранговых блоков производятся уже с учетом измененного СКО всего изображения.

Предположим, что все ранговые блоки изображения были рассмотрены, и для некоторых из них было принято решение о разделении на подблоки. После этого те ранговые блоки, которые не подлежат дальнейшему разделению, выводятся из рассмотрения, а подблоки подлежащих разделению блоков рассматриваются как самостоятельные ранговые блоки. Таким образом, процесс рекурсивно повторяется до тех пор, пока не будет достигнуто минимальное допустимое значение размера рангового блока, либо пока любая попытка дополнительного разбиения блока не станет приводить к увеличению длины описания.

Сделаем отдельное замечание по поводу размеров доменных блоков. В предлагаемом алгоритме разбиения размер доменов остается постоянным для любого уровня квадродерева, меняется лишь его масштабирующий коэффициент. В действительности есть возможность для каждого множества ранговых блоков генерировать множество доменов так, чтобы соотношение размеров было постоянным – в общем случае это позволит достигать меньшего значения СКО, соответствующего ранговым блокам, и, как следствие, лучшего качества восстановленного изображения. Однако стоит учитывать, что одновременно с увеличением количества классов рассматриваемых доменов возрастет длина описания сжатого изображения, а также, в связи со значительным увеличением количества самих рассматриваемых доменов, существенно возрастут временные затраты на компрессию, что является нежелательным.

Экспериментальная проверка

Описанная модификация метода выбора разбиения изображения в форме квадродерева на основе принципа МДО при фрактальном сжатии была протестирована на выборке из 40 изображений различных классов. На рисунках приведены примеры разбиения некоторых изображений выборки. На рис. 1 изображения разбиты на ранговые блоки по традиционной схеме квадродерева для различных значений порога на СКО. Видно, что форма разбиения зависит от устанавливаемого порога; при этом невозможно заранее предугадать, какой именно порог обеспечит наилучшее качество сжатия. На рис. 2 приведены примеры разбиения с применением критерия МДО. В силу того, что оценка результирующего качества сжатия производится уже в процессе компрессии, можно утверждать, что полученные разбиения являются оптимальными в смысле минимизации длины описания сжатого изображения. Также отметим, что в отличие от традиционного алгоритма формирования квадродерева, включающего разбиение блоков на

подблоки по критерию порогового значения СКО, построение данных разбиений производилось без ручного задания каких-либо параметров, т.е. полностью автоматически.

В таблице приведено сравнение результатов компрессии изображений для традиционного и оптимизированного метода формирования квадродерева. В таблице указаны значения длин описания, которые характеризуют общее качество компрессии (длину сжатого файла и объем информационных потерь). Были рассмотрены три значения порога на СКО: 5, 10 и 20. Для каждого из этих значений была проведена компрессия и выбран лучший результат (однако для изображения Cloud результат при минимальном пороге оказался слишком плохим – 145208 бит, поэтому расчеты были проведены вплоть до порога на СКО, равного 3).

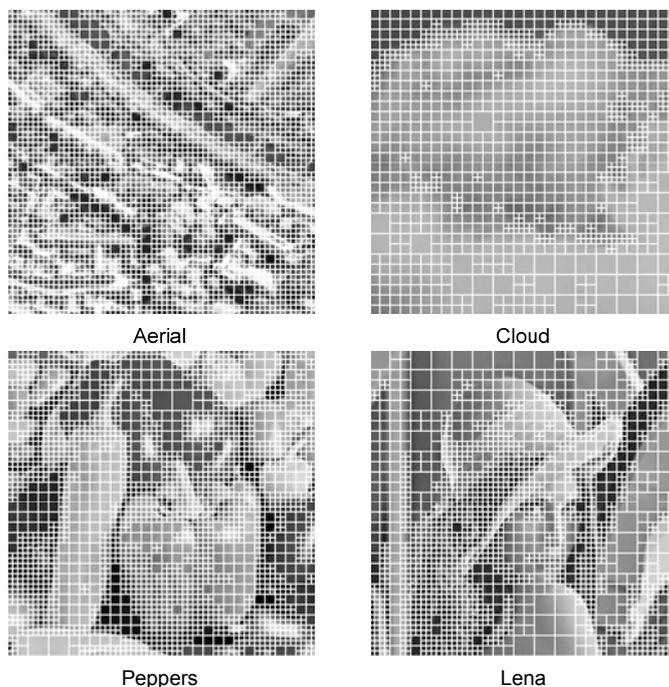


Рис. 1. Примеры разбиения в форме квадродерева с применением критерия порогового значения СКО

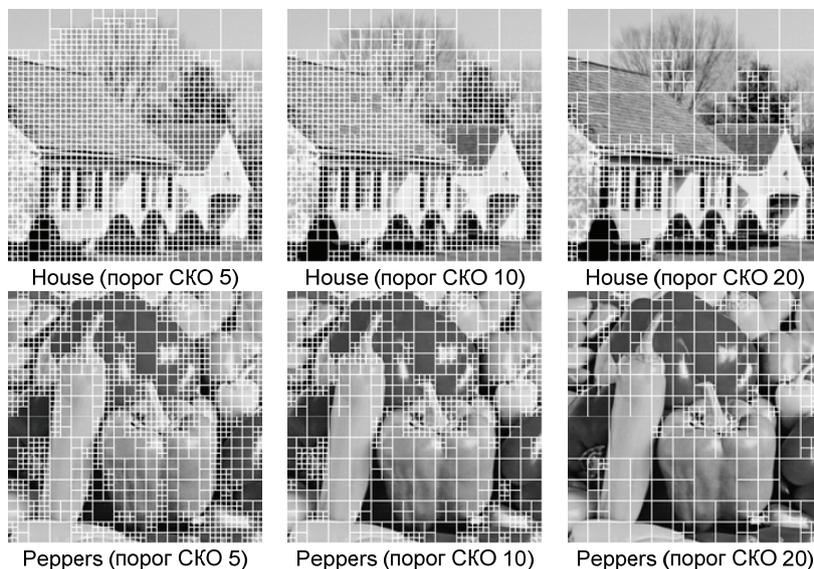


Рис. 2. Примеры разбиения в форме квадродерева с применением критерия МДО

Изображение	Порог СКО	$L_{СКО}$, бит	$L_{МДО}$, бит
Cloud	3	88530	66765
House	20	356298	340717
Lena	10	268829	248291
Peppers	10	260500	234252

Таблица. Результаты компрессии

Значения результирующих длин описания при неоптимальных порогах на СКО оказываются хуже приведенных в таблице – например, для изображения House они составляют 366519 бит (для порога 5) и 362298 бит (для порога 10).

Из таблицы видно, что выигрыш в качестве компрессии при использовании разбиения по критерию МДО составляет в среднем почти 15%. Заметим, что полученный выигрыш в длине описания эквивалентен выигрышу в 1,2–1,5 раза по степени сжатия при равном качестве восстановленного изображения или в 1,1–1,3 раза по качеству изображения (измеряемому по показателю PSNR) при равной степени компрессии.

Заключение

В работе исследован метод оптимизации разбиения в форме квадродерева для фрактального сжатия изображений. В целях оценки влияния параметров разбиения на качество сжатия предложено использовать критерий МДО. Показано, что данный критерий является адаптивным к особенностям изображений, а также учитывает не только качество восстановленного изображения, но и результирующую степень компрессии.

Хотя принцип МДО не принято применять в задачах компрессии, поскольку он сам сводится к использованию степени компрессии как критерия выбора модели, тем не менее, данный принцип оказывается здесь полезным. Это связано с тем, что при компрессии нередко выбор модели оказывается подзадачей, решаемой, однако, независимо на основе неоптимальных эвристических критериев. Успешность применения принципа МДО для улучшения метода фрактального сжатия показывает перспективность его внедрения и в другие методы компрессии, что должно улучшить их адаптивность к содержанию изображений и повысить качество сжатия.

Работа выполнена при поддержке Министерства образования и науки и Совета по грантам Президента РФ (грант МД-2040.2010.9).

Литература

1. Ghim Hwee Ong, Chorng Meng Chew, Yi Cao. A simple partitioning approach to fractal image compression // Proc. ACM Symposium on Applied Computing. – 2001. – P. 301–305.
2. Yigang Wang, Yiwen Jin, Qunsheng Peng. Merged quadtree fractal image compression // Optical Engineering. – 1998. – V. 37. – № 8. – P. 2284–2289.
3. Rissanen J. Hypothesis selection and testing by the MDL principle // The Computer Journal. – 1999. – V. 42. – № 4. – P. 260–269.
4. Vitanyi P.M.B., Li M. Minimum description length induction, Bayesianism, and Kolmogorov complexity // IEEE Trans. on Information Theory. – 2000. – V. 46. – № 2. – P. 446–464.
5. Окунев В.В., Потапов А.С. Анализ фрактального представления изображений по критерию репрезентационной минимальной длины описания // Труды научно-исследовательского центра Фотоники и оптоинформатики. Сб. статей / Под ред. И.П. Гурова и С.А. Козлова. – СПб: СПбГУ ИТМО, 2010. – Вып. 2. – С. 315–325.

- Окунев Вадим Вячеславович** – Санкт-Петербургский государственный университет информационных технологий, механики и оптики, мл. научный сотрудник, vadik-okunev@yandex.ru
- Потапов Алексей Сергеевич** – Санкт-Петербургский государственный университет информационных технологий, механики и оптики, доктор технических наук, профессор, pas.aicv@gmail.com