

Московский государственный университет имени М. В. Ломоносова
Механико-математический факультет

На правах рукописи

Шокуров Антон Вячеславович

**Метод адаптивного интерактивного анализа
растровых изображений изменяемой детализации**

Специальность 05.13.17 – Теоретические основы информатики

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

Москва – 2010

Работа выполнена на кафедре вычислительной математики
Механико-математического факультета
Московского государственного университета имени М. В. Ломоносова

Научный руководитель: доктор физико-математических наук,
профессор Михалев Александр Васильевич;

Официальные оппоненты: доктор физико-математических наук,
профессор Соколов Сергей Михайлович;

доктор технических наук,
профессор Назиров Равиль Равильевич.

Ведущая организация: Учреждение Российской Академии Наук Геофизический центр РАН.

Защита состоится 29 декабря 2010 г. в 16 часов 45 минут на заседании диссертационного совета Д.501.002.16 при Московском государственном университете имени М. В. Ломоносова по адресу: Российская Федерация, 119991, Москва, ГСП-1, Ленинские горы д. 1, МГУ, Главное здание, Механико-математический факультет, аудитория 14-08.

С диссертацией можно ознакомиться в библиотеке Механико-математического факультета МГУ имени М. В. Ломоносова (14 этаж, Главное здание).

Автореферат разослан 29 ноября 2010 г.

Ученый секретарь диссертационного совета,
доктор физико-математических наук,
профессор

А. А. Корнев

Общая характеристика работы

Актуальность работы. С развитием средств вычислительной техники стало возможным формирование растровых изображений большого размера и разрешения. Возрастающие требования, которые предъявляются при этом к качеству изображений, приводят к постоянному повышению размера и разрешения как самих растровых изображений, так и устройств их отображающих. Последнее обстоятельство вызвало потребность в исследовании и разработке методов кодирования соответствующей визуальной информации в виде данных, что влечет за собой необходимость разработки специальных моделей данных и новых принципов их проектирования. Кроме объективно возникающих вопросов с хранением данных появляется необходимость в удаленном анализе (в рамках клиент-серверной архитектуры) растровых изображений, например, по каналу связи низкой пропускной способностью. Заметим, что размер раstra запрашиваемых для анализа фрагментов изображений может существенно различаться в зависимости от специфики класса анализируемых изображений и от конкретных потребностей специалиста. Так, при отображении фрагмента изображения на стенде, состоящем из нескольких мониторов (или экранов), необходимо извлекать фрагменты большого размера, а при отображении на переносном устройстве либо при изучении большого количества разрозненных фрагментов появляется потребность в извлечении фрагментов небольшого размера. Данные произвольного фрагмента при этом должны передаваться последовательно, в порядке, обеспечивающем постепенное улучшение качества восстанавливаемого фрагмента вплоть до точного восстановления. Как следствие, появляется необходимость в разработке и исследовании метода адаптивного интерактивного анализа растровых изображений.

Важной характеристикой растровых изображений является *степень детализации*, определяющая выбор оптимального разрешения, при котором не будут потеряны существенные детали растрового изображения. Во многих растровых изображениях степень детализации варьируется, то есть *растровые изображения имеют области с изменяемыми степенями детализации*. Фотографии, например, содержат как попавшие, так и не попавшие в ГРИП (глубина резко изображаемого пространства) области. Соответственно, попавшие в ГРИП области обычно имеют более высокую степень детализации, чем не попавшие в нее. Использование информации о вариациях степени детализации позволило бы существенно увеличить степень сжатия алгоритма кодирования и существенно уменьшить объем извлекаемых данных соответствующего алгоритма декодирования. Существующий способ (основанный на понятии *региона интереса*) в недостаточной степени (только на стадии квантования) использует информацию об областях изменяемой детализации, что отрицательно влияет на объем кодированного потока данных и на процесс извлечения фрагмента. По-

следнее подчеркивает важность разработки метода фильтрации неактуальной визуальной информации.

С учетом представленных соображений тема диссертации является актуальной и соответствует областям исследований в сфере теоретической информатики, а именно – к исследованию и разработке методов и алгоритмов анализа, распознавания и синтеза растровых изображений, что отражено в п.п. 5, 7 Паспорта специальности 05.13.17 “Теоретические основы информатики”.

Задача интерактивного анализа растровых изображений в той или иной степени решается путем применения программных средств, реализующих метод (стандарт) JPEG, который основан на *дискретном косинусном преобразовании* (ДКП). Однако, большинство методов, основанных на ДКП имеют существенные недостатки. Альтернативный подход к кодированию растровых изображений основан на теории вейвлет-преобразований, который устраняет большинство имеющихся у ДКП недостатков, а именно – устраняется избыточность кодированного потока данных и объем передаваемой клиенту данных.

Можно выделить два способа при кодировании вейвлет-преобразованного растрового изображения, а именно — внутриподдиапазонный и межподдиапазонный. Внутриподдиапазонные методы (например, EBCOT, SPECK и SWEET) основаны на устранении корреляции между соседними вейвлет-коэффициентами в каждом из поддиапазонов по отдельности. К недостатку метода EBCOT можно отнести сложный характер его преобразования. К недостаткам самого метода JPEG2000 следует отнести существенное время, затрачиваемое на извлечение фрагментов, и несоответствие объема данных при извлечении небольших фрагментов. Межподдиапазонные методы основаны на устранении корреляции между вейвлет-коэффициентами, принадлежащими к разным вейвлет-поддиапазнам. Метод SPIHT является ключевым в данной категории и его можно использовать при произвольном вейвлет-преобразовании. Данный метод очень часто применяется при сравнительном тестировании новых методов кодирования растровых изображений.

Каждый из известных методов кодирования изображений имеет свою область применения ввиду того, что в каждой области существуют соответствующие им требования к программной реализации. В частности, метод SPIHT активно исследуется и используется применительно к задачам космического, медицинского и военного характера. Выбор метода SPIHT для решения этих задач был обусловлен в том числе и перечисленными ранее соображениями. С момента разработки метода SPIHT и по настоящее время на базе данного метода учеными разработано большое число улучшений и расширений. В частности, была разработана возможность извлечения всего растрового изображения при различных масштабах. Однако задача извлечения фрагмента растрового изображения при различных масштабах в общем случае решена не была.

Все изложенное выше позволяет сделать вывод, что проблема кодирова-

ния и адаптивного интерактивного анализа растровых изображений большого размера, а тем более – растровых изображений изменяемой детализации, еще недостаточно исследована. Принимая во внимание перечисленные ранее результаты сравнительного анализа существующих методов кодирования растровых изображений было решено при разработке адаптивного интерактивного метода анализа растровых изображений изменяемой детализации использовать принципы лежащие в основе метода SPIHT. Это определило выбор темы диссертации, которая актуальна в теоретическом, в практическом плане и посвящена исследованию методов кодирования визуальной информации, разработке математических моделей таких данных и программных механизмов управления ими.

Предметом исследования данной работы является метод кодирования, построенный на принципах, положенных в основу метода SPIHT преобразования визуальной информации в данные, допускающий адаптивный интерактивный анализ растрового изображения изменяемой детализации.

Целью диссертационной работы является разработка на базе традиционного метода SPIHT эффективного метода адаптивного интерактивного анализа растрового изображения изменяемой детализации и исследование свойств его программной реализации на тестовых изображениях. Эффективность включает объемы получаемых кодированных потоков данных, скорость извлечения фрагментов (размер растра которых может составлять от нескольких десятков до нескольких тысяч пикселей по каждому из измерений), количество соответствующих извлекаемым фрагментам данных и объем используемой оперативной памяти.

К числу новых научных положений, представленных в диссертации, можно отнести следующие.

- Представление растрового изображения изменяемой детализации, основанное на его вейвлет-представлении.
- Метод кодирования, основанный на методе SPIHT, данного представления, допускающий интерактивный анализ изображения. Доказана независимость объема используемой алгоритмами кодирования/декодирования оперативной памяти от размера данного изображения.
- Алгоритм формирования кодированного потока данных согласно кривой Гильберта, который при извлечении фрагмента растрового изображения позволяет существенно снизить (доказаны оценки) степень загрузки вычислительной системы по сравнению с построчным способом формирования потока данных.
- Доказана теорема, условия которой гарантируют, что при формировании кодированного потока данных под извлечение небольших фрагментов объем извлекаемых алгоритмом декодирования данных не увеличивается, а в невырожденном случае — уменьшается.
- Алгоритм адаптивного извлечения фрагментов изображения, который позво-

ляет существенно снизить (доказаны оценки) степень загрузки вычислительной системы при извлечении больших фрагментов из кодированного потока данных сформированного под извлечение небольших фрагментов.

- Программная реализация разработанного метода и результаты её тестовых испытаний.

Методы исследований. В работе использованы методы вейвлет-анализа, кодирования информации, топологии, компьютерной графики и системного программирования.

Теоретическая значимость работы. Применение методов на базе метода SPIHT не ограничивается исключительно двумерным пространством, использованием вейвлет-преобразования и кодированием визуальных данных. Существует большое число статей, посвященных альтернативному применению метода SPIHT, в частности, применительно к трехмерному пространству, к совместному с ДКП и EBCOT кодированию и к кодированию аудио данных. По этой причине задачу, решение которой представлено в диссертации, можно рассматривать как заявку на решение аналогичной задачи при альтернативных применениях метода SPIHT.

Практическая значимость работы. Программная реализация описываемого в диссертационной работе метода, кроме решения поставленной цели, при кодировании растровых изображений сравнима по степени сжатия и характеристики ПОСШ¹ с известными методами, которые активно используются на практике. В остальных случаях (специальные классы изображений изменяемой детализации, при извлечении небольших и больших фрагментов) предложенный метод показывает существенно лучший результат.

Достоверность научных положений. Все научные положения, сформулированные в диссертации, доказываются и подтверждаются численными экспериментами, проведенными для различных растровых изображений.

Апробация работы. Результаты работы были представлены на “Ломоносовских чтениях—2007” на кафедре вычислительной математики мех-мат факультета МГУ им. М.В. Ломоносова в 2007 г.; на специальном семинаре “Машинная графика” на мех-мат факультете МГУ им. М.В. Ломоносова в 2008 г.; на семинаре “Машинная графика и обработка изображений” на факультете ВМиК МГУ им. М.В. Ломоносова в 2008/2010 г.; на объединённом семинаре по робототехническим системам ИПМ им. М.В. Келдыша РАН, МГУ им. М.В. Ломоносова, МГТУ им. Н.Э. Баумана, ИНОТиИ РГГУ в 2010 г.; на семинаре “Проблемы современных информационно-вычислительных систем” на мех-мат факультете МГУ им. М.В. Ломоносова в 2010 г.; на совместном семинаре Геофизического центра РАН и Института космических исследований РАН в 2010 г.

¹ Пиковое отношение сигнала к шуму (в англоязычной литературе PSNR) является стандартной метрикой (измеряется в децибелах) измерения схожести между двумя растровыми изображениями.

Публикации. По теме диссертации опубликованы 3 статьи (см. [А1, А2, А3]) в журналах из перечня, рекомендованного ВАК Минобрнауки России.

Личный вклад автора. Все результаты диссертации были получены автором самостоятельно. Разработанное автором программное обеспечение на практике подтвердило соответствие предъявляемых к нему требований и выполнение описанных в работе свойств.

Структура и объем диссертации. Диссертационная работа объемом 129 страниц состоит из введения, краткого обзора истории исследований, пяти глав, заключения, списка публикаций и цитированной литературы. Работа содержит 34 рисунка, 7 таблиц и 46 алгоритмов. Список литературы включает 155 наименования. Приложения к диссертационной работе составляют 13 страниц.

Содержание работы

Во введении обоснована актуальность диссертационной работы, аргументирована научная новизна исследований, сформулирована цель, представлены новые научные положения, показана практическая и теоретическая значимость полученных результатов.

В главе 1 дано авторское определение обобщенного растрового изображения, в главе 2 представлена авторская схема его кодирования, а в главе 3, соответственно, схема извлечения фрагмента. В главе 4 автором доказана основная теорема. В главе 5 приведены сравнительные результаты практической работы программной реализации представленного в работе метода, в приложении А изложены основы теории дискретного вейвлет-анализа, а в приложении Б обоснована корректность проводимого тестового испытания путем сравнительного анализа различных программных реализаций метода JPEG2000.

Краткий **обзор истории исследований**, предшествующих работе, позволяет лучше понять её истоки и побудительные мотивы к её выполнению. Задача интерактивного анализа растровых изображений в той или иной степени решается путем применения программных средств, реализующих метод (стандарт) JPEG, который основан на *дискретном косинусном преобразовании* (ДКП). Однако, большинство методов, основанных на ДКП, имеют существенные недостатки. Во-первых, ДКП фиксировано и, как следствие, не может быть адаптировано в зависимости от входного растрового изображения. Во-вторых, многомасштабность, необходимая при интерактивном анализе, добавляется искусственным образом. Этот факт означает, что по растровому изображению вначале строится пирамида его масштабов (методом отличным от ДКП), а потом к каждому уровню пирамиды применяется метод, основанный на ДКП. Кроме того, при больших степенях сжатия визуально заметна блочность кодирования. Основным преимуществом существующих методов, основанных на методе ДКП, является их скорость кодирования/декодирования. Отметим, что реше-

ние задачи интерактивного анализа достигается только за счет использования существенно большего по объему кодированного потока данных.

Альтернативный подход к кодированию растровых изображений основан на теории вейвлет-преобразований. К преимуществам данного подхода можно отнести возможность адаптирования вейвлет-преобразования под входное изображение, что позволяет существенно повысить степень сжатия методов кодирования для определенных (специальных) классов растровых изображений. К тому же, масштабируемость достигается естественным образом, так как вейвлет-преобразование, используемое для декорреляции данных растрового изображения, автоматически строит и его масштабы. Благодаря свойству компактизации энергии и её более точному соответствию зрительной системе человека, большинство методов, основанных на вейвлет-преобразовании, превосходят метод JPEG по качеству восстановления растрового изображения как объективно (согласно метрике ПОСШ), так и субъективно (согласно средней оценке мнения группы экспертов). Следует также отметить, что при использовании небольшой начальной части потока данных, методам, основанным на вейвлет-преобразовании, присущи более благоприятные визуальные погрешности в сравнении с методом JPEG. В частности, алгоритмы распознавания образов (лица, отпечатки пальцев и тому подобные) в среднем работают лучше на изображениях, которые были закодированы с использованием методов SPIHT и JPEG2000, основанные на вейвлет-преобразовании, чем закодированные с использованием метода JPEG.

Можно выделить два способа при кодировании вейвлет-преобразованного растрового изображения, а именно — внутривидиапазонный и межвидиапазонный.

Внутривидиапазонные методы (например, EBCOT, SPECK и SWEET) основаны на устранении корреляции между соседними вейвлет-коэффициентами в каждом из поддиапазонов по отдельности. Заметим, что метод EBCOT был взят за основу стандарта JPEG2000, который также в той или иной степени решает задачу интерактивного анализа изображений. К недостатку метода EBCOT можно отнести сложный и, как следствие, фиксированный характер его преобразования под кодирование растровых изображений, что означает отсутствие возможности его существенной модификации. К недостаткам самого метода JPEG2000 можно отнести существенное время, затрачиваемое на извлечение фрагментов, и несоответствие объема извлеченных данных размеру раstra небольших фрагментов.

Межвидиапазонные методы основаны на устранении корреляции между вейвлет-коэффициентами, принадлежащими к разным вейвлет-поддиапазнам. Ключевым понятием в таких методах является *нуль-дерево*, в основе которого лежит гипотеза о том, что если вейвлет-коэффициент меньше определенного порога, то скорее всего и его *дети* (на вейвлет-преобразовании можно есте-

ственным образом задать иерархическую древовидную структуру) тоже меньше. Метод SPIHT (который можно использовать при произвольном вейвлет-преобразовании) является ключевым в данной категории и часто используется при сравнительном тестировании новых методов кодирования изображений.

Следует отметить, что методы кодирования текстур также позволяют интерактивно анализировать изображения (определяющим свойством методов кодирования текстур как раз и является возможность случайного доступа к любой его части при произвольном масштабе). Однако, например, они не обладают (так как это не требуется) возможностью передачи данных с постепенным повышением качества восстанавливаемого изображения (согласно метрики ПОСШ).

Каждый из перечисленных методов имеет свою область применения ввиду того, что в каждой области существуют соответствующие им требования к программной реализации (например, скорость работы, используемый объем оперативной и дисковой памяти и качество восстанавливаемого изображения). Более того, со временем требования к качеству кодирования/декодирования решаемых задач только повышаются. В зависимости от области применения (для повышения эффективности) специфицируется также и класс кодируемых изображений. Так, метод SPIHT активно исследуется и финансируется применительно к задачам космического, медицинского и военного характера. Выбор метода SPIHT для решения этих задач был обусловлен в том числе и ранее перечисленными соображениями.

С момента разработки метода SPIHT и по настоящее время на базе данного метода учеными было разработано большое число улучшений и расширений. Эти исследования касаются повышения качества восстановления изображения и степени сжатия, устойчивости к ошибкам в канале связи, упрощениям и реализации метода SPIHT на специализированных вычислительных машинах. В частности, исследователями была разработана возможность извлечения всего растрового изображения при различных масштабах. Однако необходимо отметить, что задача извлечения фрагмента растрового изображения при различных масштабах в общем случае решена не была.

В **главе 1** дано определение растрового изображения и введено авторское определение обобщенного растрового изображения.

В **разделе 1.1** в случае кодирования изображений большого размера обоснована несодержательность понятия растрового изображения. Затем, учитывая представленные недостатки растрового изображения, формализовано понятие растрового изображения с областями изменяемой детализации. Последний факт означает, что введено понятие фильтра τ , который для каждой области задает её степень детализации $detail$ и положение, заданное в координатной системе растрового изображения. Причем, степень детализации равная 0 соответствует самому высокому разрешению, а с увеличением числа $detail$ разрешение области понижается в 2^{detail} раз.

Вейвлет-представление, будучи многомасштабным, естественным образом использовано для задания на нем структуры, учитывающей степени детализации областей. Учитывая это обстоятельство, обозначим через \tilde{C}_N матрицу N -каскадного вейвлет-преобразования изображения I , где $N := \max_{i=1}^M \text{detail}_i$. В работе предложено использовать фильтр τ для *фильтрации* вейвлет-коэффициентов. Последний факт означает, что те вейвлет-коэффициенты, которым не соответствуют области необходимой степени детализации, полагаются равными ∞ . Вейвлет-коэффициенты не равные ∞ считаются *активными* и только они кодируются.

Определение 1. *Обобщенное растровое изображение*, отвечающее растровому изображению I и числу $N \in \mathbb{N}$, есть пара: $\Theta := (N, \tilde{C}_N)$.

В разделе 1.2 приведены примеры задания обобщенного изображения. На рис. 1 показан пример растрового изображения с областями изменяемой детализации. Фильтр τ для данного изображения задан как

$$\tau := \{512, 512, \mathcal{R}egion_{\text{лицо}}, \mathcal{R}egion_{\text{фон}}\},$$

где область $\mathcal{R}egion_{\text{лицо}}$ соответствует лицу и имеет степень детализации равной 0, а область $\mathcal{R}egion_{\text{фон}}$ задает фон и имеет степень детализации, например, равную 3.



Рис. 1. Пример фотографии с областями изменяемой детализации. Топология τ состоит из двух областей, разделенных ломаной, которая задана как последовательность вершин: $(390, 0)$, $(390, 310)$, $(412, 330)$, $(427, 410)$, $(395, 410)$, $(352, 380)$, $(277, 470)$, $(170, 480)$, $(112, 400)$, $(110, 290)$, $(58, 220)$.

Заметим, что фильтр τ считается известным априори как серверу, так и клиенту. Поэтому фильтр не кодируется и не передается при запросе фрагмента изображения. Это является ключевым при разработке нового объекта — зачем кодировать и передавать то, что известно из каких-то других соображений. Вторая глава посвящена кодированию вейвлет-коэффициентов, считая что фильтр τ известен.

В разделе 1.3 приведены выводы, а именно – понятие растровое изображение несодержательно на практике. По этой причине автором разработана новая модель, которая учитывает заданные пользователем степени детализации областей в изображении.

В главе 2 описана авторская схема (набор взаимосвязанных алгоритмов) кодирования обобщенного растрового изображения. Схема кодирования состоит из первичного кодирования обобщенного растрового изображения, вычисления дискретного вейвлет-преобразования обобщенного растрового изображения, кодирования соответствующих вейвлет-коэффициентов, сжатия пакетов данных и формирования целевого потока данных.

В разделе 2.1 представлены исходные положения построения схем кодирования/декодирования. В частности, продемонстрирован механизм построения алгоритма, позволяющий оценить объем оперативной памяти, используемый соответствующей программной реализацией.

В разделе 2.2 показан *первичный* способ кодирования обобщенного растрового изображения. Целью разработки первичного способа кодирования является повышение эффективности разрабатываемой системы кодирования. В частности, для повышения корреляции между соседними вейвлет-коэффициентами, решено разбить поддиапазоны N -каскадного вейвлет-представления на блоки небольшого размера (например, 64×64 вейвлет-коэффициентов), что улучшит кешируемость соседних вейвлет-коэффициентов. Для обработки всех вейвлет-коэффициентов при этом достаточно обработать вейвлет-коэффициенты всех блоков. Представленный подход позволяет разработать такие алгоритмы, что используемый их программными реализациями объем оперативной памяти не зависел бы от размера изображения.

В работе автором также предложен новый порядок хранения данных, соответствующих блокам, в потоке. Заметим, что в известных методах кодирования растровых изображений (например, в JPEG2000 и JPEG) порядок хранения блоков построчный. Такой порядок неэффективен при извлечении прямоугольных фрагментов растрового изображения. В новом подходе данные, соответствующие блокам, предложено хранить согласно кривой Гильберта, что влечет равноценную корреляцию между блоками как по горизонтали, так и по вертикали. Последний факт означает, что данные соседних блоков эффективнее кешируются. Предложенный порядок хранения блоков в сравнении с построчным порядком позволяет существенно сократить степень загрузки системы при из-

влечении фрагментов растрового изображения (обоснование дано в главе 3). Предложенные механизмы также актуальны при разработке алгоритмов, которые обрабатывают (например, вычисляют вейвлет-преобразование или кодируют) первичное представление изображения.

В разделе 2.3 автором представлено расширение дискретного вейвлет-преобразования на случай обобщенного растрового изображения и приведен алгоритм его вычисления.

Вейвлет-преобразование обобщенного растрового изображения Θ определено как соответствующее вейвлет-преобразование поддиапазона LL (в работе показано, что поддиапазон LL обобщенного растрового изображения Θ состоит только из активных узлов, по этой причине он суть растровое изображение). Из определения следует, что вейвлет-преобразование обобщенного растрового изображения есть суть обобщенное растровое изображение. В частности, N -каскадное вейвлет-преобразование растрового изображения есть обобщенное растровое изображение. По этой причине, при представлении схемы кодирования считается, что все необходимые (в рамках метода кодирования изображений) вейвлет-преобразования к изображению уже применены.

В работе автором предложен алгоритм вычисления дискретного вейвлет-преобразования, который основан на “оконном” подходе. Последний факт означает, что сначала вычисляется дискретное вейвлет-преобразование (поддиапазоны LL, LH, HL и HH) некой окрестности (зависящей от типа вейвлет-преобразования) четверки блоков (2×2) исходного растрового изображения. Затем, из соответствующих вычисленных поддиапазонов LL, LH, HL и HH вырезаются блоки, которые являются искомыми блоками соответствующих поддиапазонов вейвлет-представления растрового изображения. Дискретное вейвлет-преобразование всего растрового изображения вычисляется путем обеспечения выполнения данного процесса для всех четверок блоков. Данный алгоритм использует первичное представление обобщенного растрового изображения, что позволяет ему реализовывать качественные преимущества первичного представления. Отметим, что данный алгоритм обеспечивает порядок вычисления блоков поддиапазонов согласованный с кривой Гильберта. Последний факт означает, что как при каскадном применении данного алгоритма (именно таким образом в работе предложено вычислять N -каскадное вейвлет-преобразование), так и при кодировании вычисленных вейвлет-коэффициентов эффективность первичного представления сохраняется.

В разделе 2.4 описана третья стадия — метод кодирования вейвлет-коэффициентов обобщенного растрового изображения.

В ходе решения автором задачи кодирования обобщенного растрового изображения сначала напоминает метод SPIHT кодирования растровых изображений. Метод SPIHT основан на том, что вейвлет-коэффициенты кодируются побитно в порядке их вклада в ПОСШ характеристику. Последовательность,

согласно которой кодируются вейвлет-коэффициенты, при этом явно не кодируется, а используется процесс разбиения множества коэффициентов на подмножества.

Первоначально имеется некое разбиение множества узлов вейвлет-преобразованного растрового изображения $\Lambda_{\text{width,height}}$:

$$\Lambda_{\text{width,height}} = \bigsqcup_m \tau_m.$$

Далее выполняется проверка значимости каждого из этих подмножеств:

$$\max_{(i,j) \in \tau_m} (|c_{ij}|) \geq 2^n,$$

и если проверка даёт ложь, то это означает, что все вейвлет-коэффициенты, соответствующие узлам подмножества τ_m , незначимы. Если же проверка даёт истину, то это означает, что хотя бы какой-то узел подмножества τ_m соответствует значимому вейвлет-коэффициенту. В данном случае значимое подмножество τ_m разбивается на подмножества $\tau_{m,p}$ по заранее оговоренным правилам, и та же самая проверка выполняется для каждого из этих $\tau_{m,p}$ новых подмножеств. Такое разбиение выполняется до тех пор, пока не будут выявлены все одноэлементные значимые множества, и, тем самым, пока не будут известны все значимые вейвлет-коэффициенты.

Чтобы уменьшить количество сравнений при сортировке и, тем самым, количество выводимой в поток битов, методы кодирования используют такие правила разбиения на подмножества, при которых незначимые подмножества содержали бы большое количество узлов, а значимые подмножества были бы одноэлементными. В методе SPIHT правила разбиения основаны на понятии нуль-дерева, а именно – подмножества представляют собой поддеревья (в том числе и усеченные) вейвлет-преобразования. Если предположить, что гипотеза об нуль-дереве верна, то поддеревья остаются неразбитыми достаточно долго.

В традиционном методе SPIHT используются три глобальных списка для хранения множеств различных типов (для каждого из которых существуют свои правила разбиения): *список незначимых множеств* (\mathcal{L}_{LIS} – list of insignificant sets), *список незначимых пикселей* (\mathcal{L}_{LIP} – list of insignificant pixels), *список значимых пикселей* (\mathcal{L}_{LSP} – list of significant pixels). Данные списки обрабатываются схемами кодирования/декодирования в одинаковом порядке.

Далее автором решается задача пространственно-многомасштабного кодирования растровых изображений (обобщенных растровых изображений все вейвлет-коэффициенты которых активны). В работе представлен новый метод кодирования растровых изображений SS-SPIHT (Spatially Scalable SPIHT), концептуально основанный на методе SPIHT. В работе предложен метод SS-SPIHT, который основан на том, что в нем локализуются глобальные списки метода

SPIHT. Последний факт означает, что для каждого из блоков вейвлет-преобразования вводятся свои три локальные списка, а именно – $\mathcal{L}_{LIP}^{l,rc}$, $\mathcal{L}_{LIS}^{l,rc}$ и $\mathcal{L}_{LSP}^{l,rc}$, где l это номер вейвлет-уровня, а rc индекс блока в данном уровне.

Помимо данной модификации, автором значительно изменен алгоритм добавления элемента в список. В случае метода SPIHT элементы добавляются в глобальные списки. В отличие от традиционного алгоритма в работе автором предложен алгоритм, который сначала вычисляет индекс $(\acute{l}, \acute{k}\acute{r})$ того блока в список которого необходимо добавить элемент. Далее рассматриваемый алгоритмом элемент добавляется в список того блока, индекс $(\acute{l}, \acute{k}\acute{r})$ которого ранее был вычислен.

Предложенный автором метод SS-SPIHT позволяет локализовать кодированные данные (биты) относительно масштаба и положения. Последний факт означает, что биты, выводимые при обработке каждого из блоков, обособляются в *пакеты данных*. Кодированный поток данных (метода SPIHT) фактически разбивается на пакеты данных. Благодаря разбиению потока на пакеты, представленный (далее) метод декодирования SS-SPIHT позволяет восстановить произвольные фрагменты растрового изображения при требуемых масштабах. Последнее достигается путем извлечения пакетов данных (из кодированного потока), которых достаточно для декодирования требуемого фрагмента.

Далее в этом разделе описывается авторское расширение GI-SS-SPIHT (General Image SS-SPIHT) метода SS-SPIHT на случай обобщенных растровых изображений. Отличительной особенностью расширенного алгоритма является то, что при кодировании учитывается активность узлов (соответствующих вейвлет-коэффициентов). Последний факт означает, что кодируются только активные вейвлет-коэффициенты, в то время как известные методы (например, подход основанный на понятие региона интереса) кодируют все вейвлет-коэффициенты (в том числе и неактивные). Заметим, что в работе множество активных узлов ActiveGrid строится таким образом, что если родитель неактивен, то и его дети неактивны. Такой подход гарантирует корректность работы метода GI-SS-SPIHT, а именно – все активные вейвлет-коэффициенты будут обработаны (закодированы).

Помимо указанных модификаций структура метода GI-SS-SPIHT также претерпевает существенное изменение: блоки обрабатываются (кодируются соответствующие вейвлет-коэффициенты) согласно кривой Гильберта, что согласуется с порядком хранения блоков вейвлет-коэффициентов в первичном представлении обобщенного изображения. В результате кодирования формируются пакеты данных, которые также хранятся в потоке согласно кривой Гильберта. Более того, в методе используется согласованность кривых Гильберта с разных вейвлет-поддиапазонов, а именно – после обработки блока-родителя обрабатываются блоки-дети (в соответствующем следующем поддиапазоне). Представленный в работе подход позволяет программной реализации метода GI-SS-

SPIHT использовать объем оперативной памяти, который не зависит от размера обобщенного растрового изображения (но зависит, например, от количества вейвлет-уровней).

В разделе 2.5 сначала дополнены исходные положения. В частности, введено понятие байтовый (например, 8-битный) поток, а на его основе определено понятие *степень сжатия* применительно к алгоритму кодирования.

Далее в разделе представлен простейший способ повышения степени сжатия пакетов данных, который заключается в уплотнении их хранения. При использовании байтового потока пакеты необходимо выравнивать на байтовые границы, что влечет потерю неопределенных битов, которые расположены в последнем байте каждого пакета. В работе решено при формировании очередного пакета (при очередном проходе), относящегося к данному блоку, записывать биты (недостающие до байта) в последний байт текущего пакета.

Затем, представлена авторская вероятностная модель (процесса вывода битов) метода GI-SS-SPIHT, которая основана на том, что при инициализации вероятностной модели текущего блока используется соответствующая модель блока-родителя. Благодаря представленной модели сохраняется эффективность использования метода арифметического кодирования даже при использовании блоков маленького размера (например, 8×8 пикселей). Заметим, что с уменьшением размера блока у известных алгоритмов (например, у алгоритма JPEG2000) существенно снижается степень сжатия, причем в наибольшей степени из-за снижения эффективности метода арифметического кодирования.

В данном разделе далее предложен новый механизм завершения прохода при кодировании очередного пакета данных. Положения, лежащие в основе алгоритма арифметического кодирования, при декодировании влекут необходимость опережающего чтения данных из потока. Последний факт означает, что при декодировании пакета необходим механизм, препятствующий чтению байтов из других (следующих за текущим) пакетов. По этой причине в известных алгоритмах для определения момента завершения декодирования используется длина пакета данных, которая хранится вместе с пакетом данных. В работе автором предложен механизм, который не опирается на хранение длины пакета, а основан на определении состояния (значения внутренних переменных алгоритма) алгоритма декодирования по текущему состоянию алгоритма кодирования. Последнее означает, что в процессе кодирования вычисляется тот объем данных, который алгоритм декодирования считывает в опережающем режиме. Вычисленное количество байтов добавляется в конец каждого пакета, чтобы обеспечить корректную работы алгоритма декодирования. Эффективность представленного подхода актуальна при сравнении объема передаваемых клиенту данных, а именно – в отличие от других алгоритмов (например, JPEG2000) информация о длине пакетов клиенту не передается, что уменьшает общий объем передаваемых данных. Отметим, что информация о длине пакетов необхо-

дима только при формировании ответа на запрос клиента.

В разделе 2.6 описан авторский способ формирования целевого кодированного потока данных. Пакеты данных переупорядочиваются таким образом, чтобы они хранились согласно следованию кривой Гильберта (отдельно для каждого уровня l и битовой плоскости n).

Обработка пакетов согласно кривой Гильберта осуществляется путем рекурсивного спуска по соответствующим квадрантам данного вейвлет-уровня. Построим quadro-дерево $\Upsilon_{n,l}$, соответствующее сумме длин пакетов данных, относящихся к обрабатываемым квадрантам. Тогда корневой узел quadro-дерева $\Upsilon_{n,l}$ будет равен сумме длин всех пакетов битовой плоскости n и вейвлет-уровня l , а его дети будут равны сумме длин пакетов соответствующих квадрантов. Заметим, что в quadro-дереве достаточно сохранить длину трех из четырех квадрантов, так как сумма длин пакетов, которые соответствуют последнему квадранту, вычисляется как сумма длин всех пакетов данных (значение родителя) за вычетом суммы длин пакетов, которые соответствуют трем первым квадрантам. По дереву $\Upsilon_{n,l}$ строится вектор длин, который и кодируется.

Далее дан алгоритм кодирования длин (*вложенных*) пакетов, который совпадает с алгоритмом, используемом в методе JPEG2000.

В разделе 2.7 дан вывод, а именно – в главе 2 представлена схема кодирования обобщенного изображения. Схема кодирования формирует целевой поток данных для его эффективного использования соответствующей схемой декодирования.

В главе 3 представлен способ декодирования фрагмента изображения.

В разделе 3.1 автором показано, как вычисляются *достаточные* для восстановления фрагмента узлы вейвлет-представления. Достаточность узлов означает, что по соответствующим им вейвлет-коэффициентам можно восстановить требуемый фрагмент растрового изображения. В работе представлены формулы, которые по координатам запрашиваемого фрагмента вычисляют множество достаточных узлов. Эти формулы основаны на том, что используемое вейвлет-преобразование является локальным преобразованием.

В разделе 3.2 представлена схема декодирования фрагмента изображения: алгоритм извлечения данных из потока и алгоритм восстановления фрагмента по этим данным.

Отметим, что на запрос пользователя передаются не сами вейвлет-коэффициенты, а пакеты данных (см. рис. 2) соответствующие блокам, которым эти вейвлет-коэффициенты принадлежат.

Серверный алгоритм использует тот факт, что пакеты данных в целевом потоке хранятся согласно кривой Гильберта. Последний факт позволяет извлекать данные, соответствующие фрагменту, путем рекурсивного спуска по соответствующим квадрантам вейвлет-уровня. В частности, если среди пакетов данных, которые соответствуют квадранту, не содержится ни одного необходимого

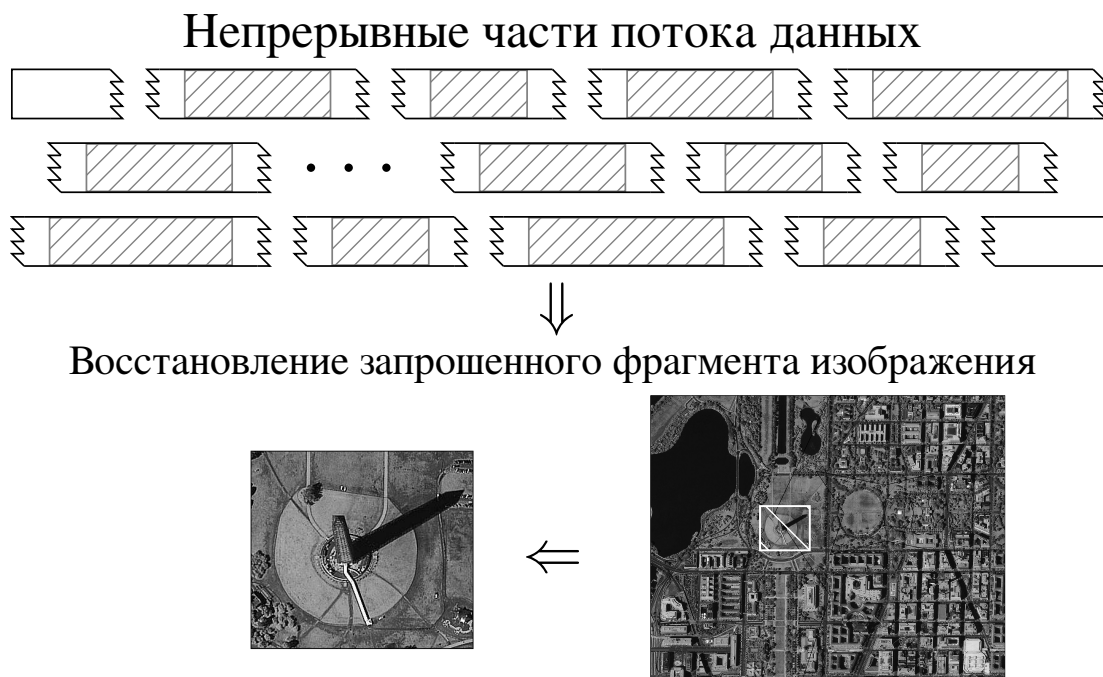


Рис. 2. На верхней части рисунка показаны непрерывные части (соответствующие последовательно идущим пакетам данных) кодированного потока по которым на нижней части рисунка изображен восстановленный фрагмент растрового изображения.

клиенту пакета, то рекурсивного спуска не производится. Аналогично, если все пакеты данных клиенту необходимы, то все данные, соответствующие квадранту, передаются (для этого используется $\Upsilon_{n,l}$) клиенту целиком и рекурсивного спуска не производится. Рекурсивный спуск производится только в том случае, когда квадранту соответствуют пакеты как отмеченные, так и не отмеченные на передачу. Представленный способ хранения пакетов позволяет существенно сократить количество обрабатываемых пакетов при извлечении данных, относящихся к фрагменту растрового изображения, что позволяет существенно снизить нагрузку вычислителя. В работе показано, что вычислительная сложность извлечения равна $O(\frac{w}{B} * \log(\frac{D}{B}))$, где $D = \max(\text{width}, \text{height})$ – линейный размер изображения, B – размер блока используемого при кодировании и w – ширина вырезаемого квадратного фрагмента.

В работе автором далее представлен алгоритм активной адаптации к размеру раstra извлекаемого фрагмента. При извлечении фрагмента размера 1000×1000 нет смысла в ответ клиенту пересылать пакеты, относящиеся к блокам размера 16×16 . В данном случае целесообразнее дать ответ пакетами, соответствующим блокам (квадрантам) размера 64×64 (что естественно увеличит общий объем передаваемых данных). При вырезании фрагмента растрового изображения, во-первых, удастся сократить количество рекурсивных вызовов, что уменьшает степень загрузки вычислительной системы, и, во-вторых, пакеты данных удастся объединять в непрерывные фрагменты кодированного потока, что уменьшает общее количество сессий доступа к хранилищу данных. Все вме-

сте это снижает загрузку вычислительной системы (стоит отметить, что связь степени адаптации с реальным повышением скорости системозависимо).

В работе автором представлены механизмы, которые играют существенную роль в уменьшении времени извлечения фрагмента растрового изображения. Отметим, что за счет использования хранилищем данных кэша (кэшируются близко расположенные данные, в частности ветви кривой Гильберта) адаптация используется неявно.

Затем в работе кратко изложен клиентский алгоритм восстановления фрагмента растрового изображения по извлеченным пакетам. Данный алгоритм строится обращением алгоритма кодирования. Отметим, что объем, используемый программной реализацией данного алгоритма оперативной памяти, зависит только от размера раstra и масштаба фрагмента, а не от размера растрового изображения.

В разделе 3.3 даны выводы, а именно – в данной главе автором представлен эффективный способ извлечения фрагментов. Критерий выбора размера блока, используемого при кодировании, рассмотрен в следующей главе.

В главе 4 автором доказана теорема, условия которой показывают эффективность представленного в работе метода. Показатель эффективности метода GI-SS-SPIHT кодирования обобщенных растровых изображений (в частности, растровых изображений) заключается в том, что объем извлекаемых данных, соответствующих восстанавливаемым фрагментам изображения, с уменьшением размера блока при кодировании не увеличивается.

В данной главе предполагается, что в схемах кодирования/декодирования алгоритм арифметического кодирования не используется. Данное условие не является недостатком, так как на практике алгоритм арифметического кодирования, в силу его большой вычислительной сложности, использовать не принято.

В данной главе под *значимыми* данными (битами) понимаются только те биты, которые были выведены в результате работы основного алгоритма кодирования. Последнее означает, что при вычислении их количества не учитываются биты, которые используются для выравнивания пакетов данных на байтовые (например, 8-битные) границы.

В разделе 4.1 автором рассмотрен вопрос инвариантности объема кодированного потока данных. Представленный в диссертационной работе метод GI-SS-SPIHT базируется на основных положениях метода SPIHT. В связи с тем, что данные методы построены на схожих положениях, метод GI-SS-SPIHT по существу переупорядочивает выводимые методом SPIHT биты.

Сначала автором доказана инвариантность объема кодированного потока данных, а именно – доказано, что количество значимых бит z , выводимых в кодированный поток данных методом GI-SS-SPIHT, не зависит от выбранного при кодировании размера блока. Затем показано, что при кодировании растрового

изображения количество значимых бит, выводимых в кодированный поток данных методом кодирования SS-SPIHT (частный случай метода GI-SS-SPIHT для случая растровых изображений), совпадает с количеством бит, выводимых в поток методом кодирования SPIHT. Более того, количество выводимых битов 1 и 0 по отдельности также будет совпадать. Для этого в работе показано, что в методе SS-SPIHT изменен только порядок обработки элементов и, следовательно, факт вывода битов сохранится.

Помимо выводимых методом GI-SS-SPIHT значимых бит, целевой кодированный поток данных содержит заголовки, длины вложенных пакетов и биты, используемые для выравнивания пакетов на естественные байтовые границы. Во втором подразделе дана оценка накладных расходов на выравнивание пакетов данных в случае массива байтов (например, из 8-чисел), а именно – оценка вклада (процентная) влияния процесса выравнивания пакетов данных на общий объем потока.

При описании механизма организации потока данных решено использовать оставшиеся биты текущего пакета данных при формировании следующего пакета, который относится к тому-же блоку. Последний факт означает, что количество битов, необходимых для выравнивания, зависит только от количества используемых блоков и не зависит от количества битовых плоскостей. По этой причине далее в работе вычисляется оценка общего количества блоков δ , на которое разбито вейвлет-преобразованное изображение. При выводе оценки считается, что $w, h \gg N$, где w и h это размер изображения, а N – количество используемых вейвлет-уровней. На практике более содержательна оценка относительной избыточности: $\epsilon\% = \frac{\delta}{wh} \lesssim \frac{1}{3s^2}$. Для блока размера 8×8 пикселей получаем оценку 0,5%, которая является незначительной. Следовательно, размер целевого потока в наибольшей степени зависит от способа хранения длин пакетов.

В разделе 4.2 сначала показывается, что количество значимых бит q , по которым можно восстановить фрагмент обобщенного растрового изображения, не зависит от выбранного при кодировании размера блока. Затем доказывается основная теорема диссертационной работы.

Напомним, что при восстановлении фрагмента не используются длины пакетов данных. В этой связи объем извлекаемых данных определяется количеством значимых данных. В работе показано, что соответствие бита фрагменту зависит только от расположения соответствующего вейвлет-коэффициента в вейвлет-преобразованном обобщенном изображении и не зависит от факта принадлежности конкретному блоку. Следовательно, общее количество выводимых значимых данных в поток, относящихся к текущему масштабу фрагмента изображения, инвариантно относительно используемого размера блока.

Далее в работе следует доказательство основной Теоремы.

Теорема 1. *С уменьшением размера блока при кодировании общий объем дан-*

ных, передаваемый клиенту на запросы фрагментов, не увеличивается, а в невырожденном случае — уменьшается.

В разделе 4.3 даны выводы, а именно – с уменьшением размера блока объемом (значимых данных) кодированного потока не увеличивается, а объем извлекаемых при этом данных (при извлечении фрагментов) может даже уменьшиться.

В главе 5 сначала (в разделе 5.1) представлена программная реализация метода GI-SS-SPIHT, а также уточнены принятые соглашения.

Далее (в разделе 5.2) приведены практические результаты работы метода GI-SS-SPIHT кодирования. Из результатов проведенных тестовых испытаний метода следует, что метод GI-SS-SPIHT имеет явное преимущество в сравнении с методом JPEG2000 по степени сжатия растровых изображений, в том числе – и с областями изменяемой детализации. Причем, с уменьшением размера блока при кодировании преимущество метода GI-SS-SPIHT становится все более существенным. В сравнении с JPEG-LS предложенный метод показывает несколько худшие результаты. Однако отметим, что метод JPEG-LS не обладает никакими другими важными свойствами (ПОСШ-характеристика и возможность извлечения фрагментов) кроме превосходящей степени сжатия.

Использование понятия обобщенного растрового изображения (см. рис. 3) действительно позволяет существенно сокращать объем данных (с контролируемой пользователем степенью качества во всем изображении), необходимых для хранения растровых изображений. Отметим, что преимущество использования обобщенного растрового изображения сохраняется в том числе и при использовании других методов кодирования (например, JPEG2000). При этом заметим, что с уменьшением размера блока при кодировании преимущество метода GI-SS-SPIHT становится все более значимым.

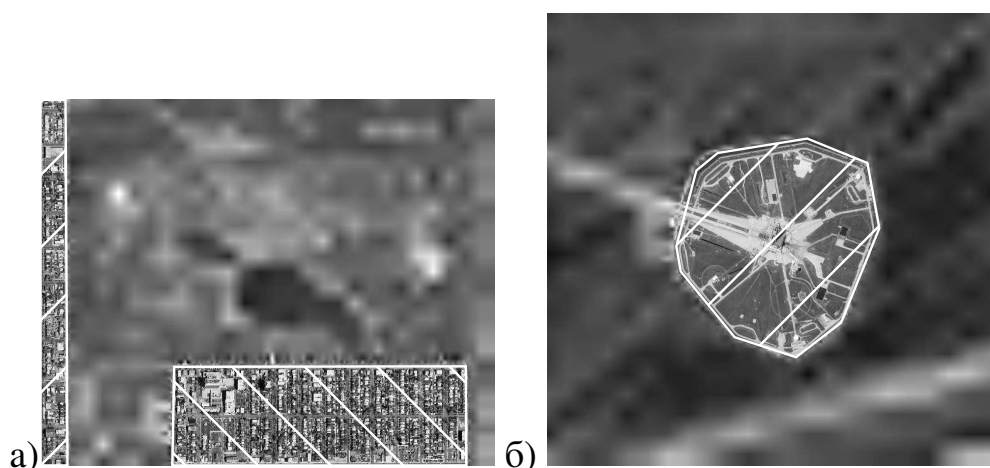


Рис. 3. Две фотографии с размеченными областями детализации: а) изображение построено по изображению зоопарка, где областями выделены группы домов; б) изображение построено по изображению космического челнока, где областью выделена взлетная площадка.

Далее приведен анализ процесса извлечения фрагмента изображения из потока. Предложенный метод GI-SS-SPIHT имеет существенное преимущество в сравнении с методом JPEG2000. Скорость извлечения фрагмента у метода GI-SS-SPIHT существенно выше (см. рис. 4), чем у метода JPEG2000, а именно – средняя скорость извлечения выше как минимум на порядок. Например, на извлечение квадратного фрагмента размера 1000×1000 пикселей предложенный метод GI-SS-SPIHT потратит 5 мс, а JPEG2000 — более 100 мс.

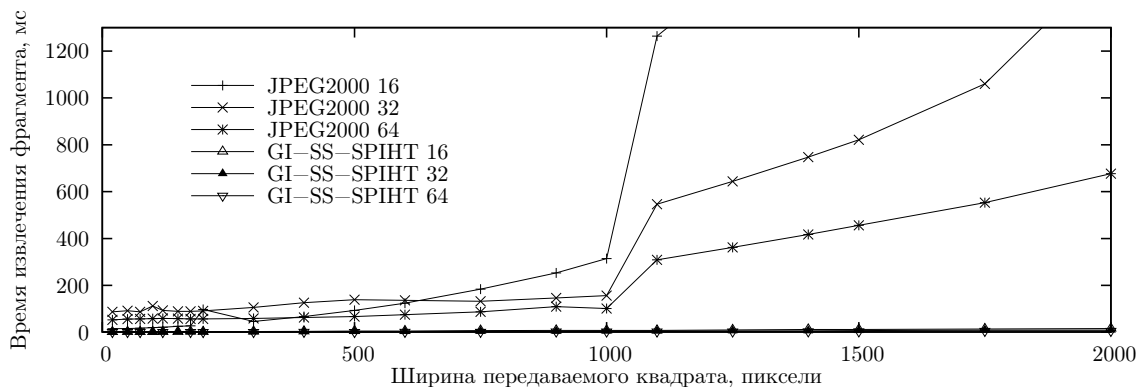


Рис. 4. Сравнение среднего времени, затраченного методами GI-SS-SPIHT и JPEG2000 на извлечение фрагмента из большого растрового изображения.

Объем извлекаемых данных методом GI-SS-SPIHT в сравнении с методом JPEG2000 в среднем ниже и он действительно уменьшается с уменьшением размера блока (согласуется с основной теоремой, притом, что используется энтропийное кодирование). Более того, представленный в настоящей работе метод GI-SS-SPIHT обладает возможностью адаптации под размер извлекаемого фрагмента. Последний факт означает, что за счет объединения соседних блоков существенно повышается скорость извлечения. Например, на извлечение фрагмента размера 1000×1000 пикселей время извлечения понизится с 8 мс. до 2 мс. Процесс восстановления фрагмента метода GI-SS-SPIHT имеет сравнимое с методом JPEG2000 качество, а использование небольших блоков заметно повышает качество. На рис. 5 приведен пример восстановления фрагмента, наглядно показывающий результат сравнения.

В настоящее время задача интерактивного анализа растровых изображений решена при определенных (существенных) ограничениях. По результатам тестовых испытаний можно сделать вывод о том, что представленный в настоящей работе метод GI-SS-SPIHT решает задачу интерактивного анализа изображений без ограничений на размеры используемых отображающих устройств и в режиме последовательной передачи данных.

К числу основных результатов, можно отнести следующие:

1. Разработан новый метод кодирования и адаптивного интерактивного анализа растрового изображения изменяемой детализации, который включает эффективные модели, алгоритмы и программные механизмы кодирования/

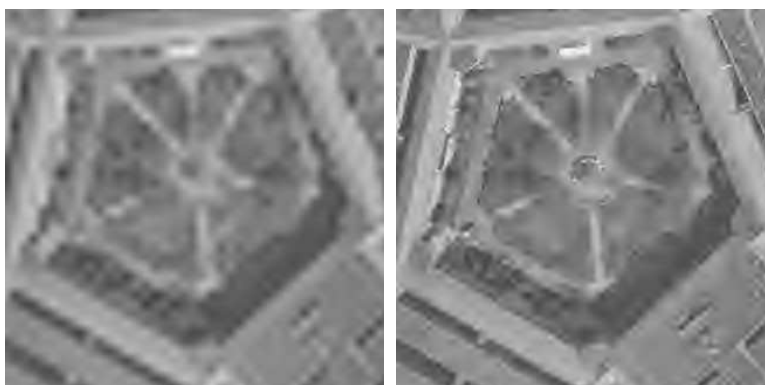


Рис. 5. Сравнение восстановленного методами JPEG2000 (слева, 25.21 дб.) и GI-SS-SPINT (справа, 26.71 дб.) соответственно фрагмента изображения 3.2.25 при 0.3 бит на пиксель. Хорошо видно, что визуальное качество правого изображения превосходит левое.

декодирования растрового изображения, использующие основные положения метода SPINT.

2. Доказательно показано, что при формировании кодированного потока данных, оптимизированного для извлечения небольших фрагментов изображения, объем данных, достаточных для их восстановления, не увеличивается.
3. Аналитические оценки алгоритмов кодирования/декодирования растрового изображения свидетельствуют о том, что объем используемой для их реализации оперативной памяти не зависит от размера этого изображения.
4. Тестовые испытания программной реализации нового метода показали, что она позволяет: последовательно извлекать из кодированного потока данные, относящиеся к определенному масштабу фрагмента изображения; адаптивно уменьшать степень загрузки вычислительной системы.

Благодарности. Автор выражает благодарность профессору А. В. Михалеву за осуществление научного руководства, профессору В. А. Васенину за ценные замечания по тексту диссертации, способствовавшие её улучшению, к.ф-м.н. старшему научному сотруднику Д. В. Иванову за внимание к работе и всем сотрудникам кафедры вычислительной математики за поддержку.

Список публикаций

- A1. Шокуров А. В. Кодирование изображений с последующим возможным оптимальным декодированием // *Фундамент. и прикл. матем.* 2007. Т. 13. С. 225–255.
- A2. Шокуров А. В., Михалев А. В. Оптимальное использование вейвлет-компонент // *Успехи матем. наук.* 2007. Т. 62. С. 171–172.
- A3. Шокуров А. В. Оптимальное использование компонент двоичного вейвлет-представления // *Вестн. Моск. ун-та.* 2009. Т. 5. С. 3–6.