

На правах рукописи

Дам Чонг Нам

**Исследование и разработка методов сжатия подвижных
изображений с использованием расширенных вейвлет-
разложений**

Специальность 05.13.01 –

Системный анализ, управление и обработка информации (информационные и
технические системы)

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата технических наук

Москва – 2019

Работа прошла апробацию на кафедре «Мультимедийные технологии и телекоммуникации» Федерального государственного автономного образовательного учреждения высшего образования «Московский физико-технический институт (национальный исследовательский университет)»

Научный руководитель:

член-корреспондент РАН,
доктор технических наук
Дворкович Александр Викторович

Ведущая организация: Ордена Трудового Красного Знамени федерального государственного бюджетного образовательного учреждения высшего образования «Московский технический университет связи и информатики»

Защита состоится 07.08.2019 в 10:00 на заседании диссертационного совета ФРКТ 05.13.01.007 по адресу 141701, Московская область, г. Долгопрудный, Институтский переулок, д. 9.

С диссертацией можно ознакомиться в библиотеке и на сайте Московского физико-технического института (национального исследовательского университета) <https://mipt.ru/education/post-graduate/soiskateli-tekhnicheskie-nauki.php>

Работа представлена «30» мая 2019 г. в Аттестационную комиссию федерального государственного автономного образовательного учреждения высшего образования «Московский физико-технический институт (национальный исследовательский университет)» для рассмотрения советом по защите диссертаций на соискание ученой степени кандидата наук, доктора наук в соответствии с п.3.1 ст. 4 Федерального закона «О науке и государственной научно-технической политике».

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы

Телекоммуникационная система включает передатчик, канал связи и приёмник, при этом канал связи должен обеспечить заданную скорость передачи информации. Чтобы уменьшить требуемую пропускную способность канала связи необходимо устранять избыточность исходных данных путем их сжатия до передачи. Чем эффективнее сжимать данные, тем меньше затрат требуется для их передачи и больше программ можно передавать через заданный канал связи с ограниченной пропускной способностью. Аппаратура обработки информации стремительно развивается, что позволяет увеличить пропускную способность канала связи. Однако требования пользователей коммуникационных услуг быстро растут, что объясняется, в том числе, появлением форматов изображений сверхвысокой четкости 4К и 8К. Так что проблема сжатия исходных данных для системы телекоммуникаций остается актуальной.

Проблема сжатия статических и подвижных изображений также является актуальной для создания центров хранения изображений и видео в цифровом виде, сформированных в медицинских целях, при космических съемках, реализованных с применением датчиков и камер систем мониторинга. Решение данной проблемы позволяет понизить объем информации, которая сохраняется на носителях.

Бурное развитие направления разработки алгоритмов и соответствующей аппаратуры цифрового сжатия статических и подвижных изображений для их передачи по каналам связи обеспечило создание ряда значимых результатов разработки алгоритмов сжатия таких изображений (широко используемые стандарты: JPEG, JPEG-2000, MPEG-1, MPEG-2, MPEG-4 (видео), H.261, H.262, H.263, H.264 (AVC), H.265 (HEVC) и др.). Эти стандарты строго определяют синтаксис цифрового потока кодированных данных и описывают самые общие

принципы его формирования. Тем не менее существует большой потенциал более эффективной обработки сигналов в рамках этих стандартов.

Кроме того, в настоящее время одной из самых быстроразвивающихся технологий является Интернет, при этом видео контент занимает огромную часть всемирного трафика. По статистике компании Cisco, в 2017 году онлайн видео составило 75% всего контента интернета, по прогнозам к 2022 году это число возрастает до 82%. Так что развитие алгоритмов сжатия видеоданных является актуальным направлением исследований.

Степень разработанности темы

Проблемой передачи и хранения статических и подвижных изображений занимаются многие исследователи в мире. К ним можно отнести: David Salomon, Khalid Sayood, Iain E.G. Richardson, D. Taubman, Nick Kingbury, Selesnick, G.J. Sullivan, и др. Среди Российских исследователей в этой области следует отметить В.П. Дворковича и А.В. Дворковича, Ю.Б. Зубарева, С.В. Умняшкина, и др.

К настоящему времени в большинстве современных видеокодеков и их реализаций, таких как HviD, VP9, h.264, Theora используется дискретное косинусное преобразование (ДКП), характерными недостатками которого является ряд эффектов, том числе, блокинг-эффект, мозаичный эффект и др. В отличие от ДКП, дискретное вейвлет-преобразование (ДВП) работает не с блоками малого размера, а с тайлами и даже с кадрами, вследствие чего ДВП лишено таких дефектов. Поэтому перспективным следует считать применение ДВП для компрессии статических и подвижных изображений. Действительно, ДВП сначала применялся в стандарте сжатия статических изображений JPEG2000. Затем появились вейвлет-видеокодеки на базе ДВП такие как Motion Wavelet, Dirac, Schrödinger, которые первоначально не получили широкого распространения из-за существенных требований к вычислительным ресурсам оборудования и значительного времени кодирования. Однако, бурное развитие

встраиваемой аппаратуры позволяет продолжать их исследовать для повышения их эффективности и усовершенствования алгоритмов сжатия подвижных изображений.

В качестве теоретической основы в 2012 году д.т.н., проф. Дворкович В.П. и д.т.н., член-корр. РАН, Дворкович А.В. предложили методику вычисления коэффициентов многоканального вейвлет-разложения, которая расширила перспективность реализации вейвлет-видеокодек. В работе *G.Y. Gryzov, A.V. Dvorkovich, "Three-Channel Wavelet Transform for Video Compression Applications", 6th Mediterranean Conference on Embedded Computing MECO 2017, 11-15 June 2017, pp. 1-4* приведены результаты внедрения трехканального вейвлет-разложения в вейвлет-видеокодек Dirac и показан существенный выигрыш трехканального вейвлет-разложения по сравнению со стандартным двухканальным вейвлет-преобразованием. В настоящей диссертационной работе основные результаты разработанных методов получены для модифицированной версии вейвлет-видеокодека Dirac после внедрения трехканального вейвлет-разложения.

Для подвижных изображений существуют три типа избыточности, такие как статистическая, внутрикадровая и межкадровая (временная). При этом с наибольшим эффектом реализуется межкадровая избыточность. При сжатии подвижных изображений ключевую роль играет процедура поиска и компенсации движения. В стандартах видеокомпрессии алгоритмы поиска и компенсации движения не фиксированы, а оставлены на усмотрение разработчиков прикладных программ. Так что разработка алгоритмов поиска и компенсации движения для того или иного видеокодека является актуальной задачей.

Нужно отметить, что для сжатия подвижных изображений дискретное вейвлет-преобразование (двухканальное или многоканальное), основанное на действительных вейвлет-коэффициентах, имеет недостаток – неинвариантность к сдвигу, заключающийся в том, что малейшие смещения изображения в кадре

приводят к значительным изменениям значений вейвлет-коэффициентов, что не позволяет использовать вейвлет-преобразования для поиска и компенсации движения. Мощным подходом к решению этой проблемы является использование комплексного вейвлет-преобразования (КВП). В диссертации рассмотрены варианты потенциального применения КВП в задаче компенсации движения для вейвлет-видеокодексов.

По перечисленным типам вейвлет-разложений, использованных в настоящей диссертационной работе, введен термин *«расширенные вейвлет-разложения»*, под которыми понимаются трехканальное вейвлет-разложение и комплексное вейвлет-разложение. К развитию методов сжатия подвижных изображений, в общем, и развитию семейства вейвлет-видеокодексов в частности в рамках проведенных исследований данной диссертационной работы предложены и реализованы оптимизированный метод квантования и способ улучшения качества традиционного метода компенсации движения для вейвлет-видеокодека Dirac.

Объектом исследования в работе являются видеокодеки, в частности вейвлет-видеокодек Dirac, для системы передачи или устройства кодирования и хранения.

Предметом исследования являются методы сжатия подвижных изображений – метод поиска и компенсации движения и метод квантования.

Целью диссертационной работы является исследование и разработка методов сжатия подвижных изображений с использованием расширенных вейвлет-разложений.

Задачи диссертационной работы

Для достижения поставленной цели в диссертации решаются следующие основные задачи:

1. исследование стандартного вейвлет-разложения;

2. исследование расширенных вейвлет-разложений – многоканального и комплексного вейвлет-разложений;
3. исследование главных функциональных блоков типичного видеокодека в общем случае и вейвлет-видеокодека Dirac в частности;
4. исследование и усовершенствование метода поиска и компенсации движения;
5. исследование и разработка оптимизированного метода квантования для вейвлет-видеокодексов;
6. внедрение разработанных методов в вейвлет-видеокодек Dirac и проведение сравнительного анализа.

Общая методика исследований

Основные результаты диссертационной работы были получены с помощью:

- ✓ теории цифровой обработки сигналов,
- ✓ теории линейных и нелинейных систем преобразования изображения,
- ✓ теории вейвлет-преобразования,
- ✓ теории вероятностей и математической статистики,
- ✓ теории информации,
- ✓ методов математического анализа,
- ✓ методов линейной алгебры,
- ✓ метода наименьших квадратов.

Для создания и отладки программного обеспечения исследований и разработок использованы язык программирования Python, системы символьной математики «Mathematica» 9, MATLAB R2013b, язык программирования C++, среда разработки Microsoft Visual Studio 2010 и 2017.

Научная новизна

К числу существенных результатов выполнения диссертационной работы, обладающих научной новизной, относятся нижеследующие.

1. Предложен и разработан метод компенсации движения с использованием комплексного вейвлет-преобразования;
2. Предложен и разработан новый способ повышения качества традиционного метода компенсации движения;
3. Предложены и разработаны некоторые подходы к оптимизации предложенного способа повышения качества традиционного метода компенсации движения;
4. Предложен и разработан оптимизированный метод квантования для модифицированного вейвлет-видеокодека Dirac.

Личный вклад автора

Автор является ответственным исполнителем во всех исследованиях и разработках по теме диссертации. Все результаты исследований получены лично автором диссертации. Первоначальные идеи и поддержка разработок принадлежат д.т.н., чл.-корр. РАН Дворковичу А.В., д.т.н., проф. Дворковичу В.П. и научному консультанту вед. инж. Грызову Г.Ю.

Основные положения, выносимые на защиту:

1. Предложенный метод компенсации движения с помощью комплексного вейвлет-преобразования позволяет снизить объем сжатого потока от 6% до 25% по сравнению с эталонными методами – методом сопоставления блоков и методом сопоставления перекрывающихся блоков, используемым в вейвлет-видеокодеке Dirac.
2. Применение предложенного способа повышения качества традиционного метода компенсации движения в вейвлет-видеокодеке Dirac может уменьшить выходной поток до 20% при PSNR = 40 дБ восстановленного

изображения (хорошее качество восстановления) и до 33% при PSNR = 42 дБ (высокое качество восстановления) для случая разбиения кадра на блоки 16×16 , от 3% до 33% при PSNR = 40 дБ и от 2% до 40% при PSNR = 42 дБ для случая разбиения кадра на блоки 32×32 . Кроме того, в сравнении с традиционным методом предложенный способ значительно удачнее работает для сложных видеопоследовательностей, в которых происходят изменение масштаба (зуммирование), вращение, значительное изменение яркости и т.п.

3. Оптимизация предложенного способа повышения качества традиционного метода компенсации движения позволяет существенно сократить объем данных по сравнению с предложенным методом до оптимизации, конкретно от 4% до 32% при PSNR = 40 дБ восстановленного изображения, от 3% до 26% при PSNR = 42 дБ для режима разбиения кадра на маленькие блоки 8×8 , до 10% при PSNR = 40 дБ, до 6,5% при PSNR = 42 дБ для случая размера блока 16×16 , до 3% при PSNR = 40 дБ, до 1,3% при PSNR = 42 дБ для случая размера блока 32×32 . По сравнению с традиционным методом предложенный способ после оптимизации может сократить выходной поток до 20% при PSNR = 40 дБ и до 33,5% при PSNR = 42 дБ для случая размера блока 16×16 , от 3% до 35% для PSNR = 40 дБ и от 2% до 40% для PSNR = 42 дБ для случая размера блок 32×32 . Необходимо также отметить, что по сравнению с положением 2 здесь увеличивается процент выигрыша до 5% для каждого тестового видео.

4. Предложенный метод квантования в вейвлет-видеокодеке Dirac позволяет заметно снизить скорость выходного потока при заданном уровне искажений. По сравнению с методом равномерного квантования, реализованного в рассматриваемом видеокодеке, предложенный метод для режима внутрикадрового кодирования уменьшает битовую скорость до 4,2% при низком уровне искажений, а для режима межкадрового кодирования уменьшает битовую скорость до 9,8% при низком уровне искажений и до 11,2% при среднем (приемлемом) уровне искажений.

Теоретическая и практическая ценность

1. Предложено и реализовано применение комплексного вейвлет-преобразования в задаче повышения точности компенсации движения;
2. Предложенные методы повышения эффективности сжатия подвижных изображений позволяют уменьшить объем кодированного потока, необходимого для хранения или передачи по каналу связи, при сохранении качества восстановленных изображений;
3. Основные результаты диссертационной работы внедрены в модифицированный вейвлет-видеокодек Dirac и в учебный процесс кафедры мультимедийных технологий и телекоммуникаций физтех-школы радиотехники и компьютерных технологий МФТИ в рамках программы дисциплины «Цифровая передача видеоинформации».

Высокая степень достоверности результатов обеспечивается применением для их получения общепринятых научных подходов и методов. Сравнения методов производились посредством численных экспериментов на классах стандартных видеопоследовательностей, рекомендованных для тестирования видеокодеков. Анализ результатов и основные выводы были сделаны на основе широко распространённой метрики качества PSNR и оценки энтропии для количества данных, необходимых для кодирования или реального значения скорости выходного потока на выходе видеокодера.

Апробация работы. Основные материалы по всем разделам диссертационной работы докладывались и обсуждались на следующих конференциях:

- 20-я Международная конференция «Цифровая обработка сигналов и ее применение», Москва, 2018 г.
- 61-я Всероссийская научная конференция МФТИ, Москва, 2018 г.

– V международная конференция Инжиниринг и Телекоммуникации - En&T, Москва, 2018 г.

Публикации. По теме диссертационного исследования опубликовано 6 научных работ, из них две статьи в рецензируемых научных изданиях, входящих в список ВАК и перечень RSCI, одна заявка на изобретение. Также принято к публикации 3 статьи, из них две статьи в рецензируемых научных изданиях, входящих в список ВАК и перечень RSCI, одна статья в изданиях, индексируемых в международных базах данных (SCOPUS, Web Of Science); 7 научных работ выполнено без соавторства.

Объем и структура диссертации. Диссертация состоит из введения, четырех глав, заключения и списка литературных источников из 94 наименований. Общий объем диссертации составляет 147 страниц, включая 69 рисунков и 27 таблиц.

Во введении обоснована актуальность темы, ее теоретическая и практическая значимости, сформулированы цель и задачи исследования, приведены направления усовершенствования методов сжатия подвижных изображения и семейства вейвлет-видеокодексов.

В первой главе даны необходимые определения и термины в кодировании изображений и видео. Приведена постановка задачи сжатия изображений и главные этапы решения этой задачи, которые подробно рассмотрены на примерах стандартов JPEG и JPEG-2000. Стандарт JPEG применяет ДКП к блокам изображения размером 8×8 пикселей. К основным недостаткам можно отнести: блокинг-эффект и мозаичный эффект, возникающие при больших степенях сжатия. Для устранения влияния таких дефектов был предложен стандарт JPEG-2000, в основе которого лежит ДВП. Этот стандарт продемонстрировал, что применение ДВП в задаче сжатия статистических и динамических изображений весьма перспективно. Действительно, в 2008 году исследовательский центр ВВС опубликовал первую версию открытого

видеокодека Dirac, основанного на ДВП. Затем также в этом году BBC продолжил опубликовать ускоренную версию видеокодека Dirac, названную видеокодеком Schrödinger. В связи с принципом работы вейвлет-видеокодеков, обрабатывающих, в отличие от стандартов H.264 и H.265, не блоки независимо, а тайлы или даже весь кадр. Это обуславливает большие вычислительные затраты, медленность обработки и трудность в распараллеливании. По этим причинам вейвлет-видеокодеки не сразу получили распространение. Однако, встраиваемая аппаратура на сегодняшний день позволяет продолжать их исследования для усовершенствования методов сжатия подвижных изображений.

В этой главе также приведен обзор методов сжатия подвижных изображения и приведены главные функциональные модули типичного кодера и декодера видео. Сделан вывод о том, что задача оптимального квантования для того или иного видеокодека является актуальной задачей в связи с тем, что наибольшие потери происходят именно на этом этапе, и оптимизированный метод квантования может значительно уменьшить объем сжатого потока при заданном качестве восстановления. Не менее актуальной является задача поиска и компенсации движения, которая занимает существенную часть вычислительных затрат при кодировании и устраняет главную проблему видеопоследовательности – временную избыточность. Далее в данной главе приведен обзор вейвлет-видеокодека Dirac, общие схемы кодера и декодера которого представлены на рисунках 1 и 2 соответственно.

Затем в настоящей главе пояснена необходимость улучшения некоторых главных модулей рассматриваемого видеокодека. Для усовершенствования методов сжатия подвижных изображений в общем и развития вейвлет-видеокодека Dirac в частности, в конце этой главы представлены некоторые задачи исследования, направленные на выполнение перечисленных актуальных задач – поиска и компенсации движения и системы квантования, которые

реализуются в блоках «анализ движения», «компенсация движения», «квантование» и «обратное квантование», представленных на рисунках 1 и 2.

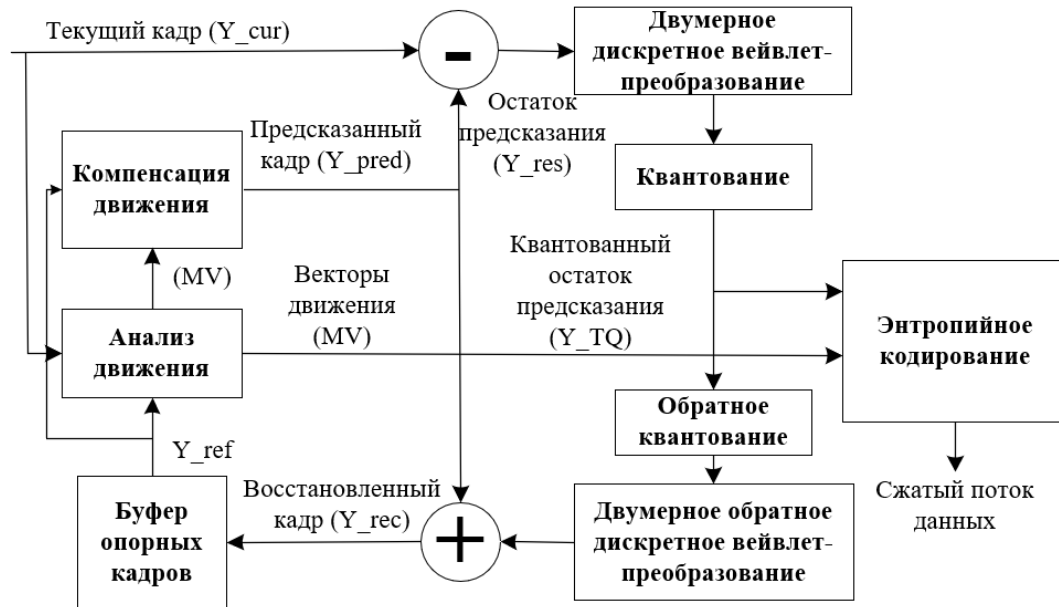


Рисунок 1. Общая схема вейвлет-видеокодера

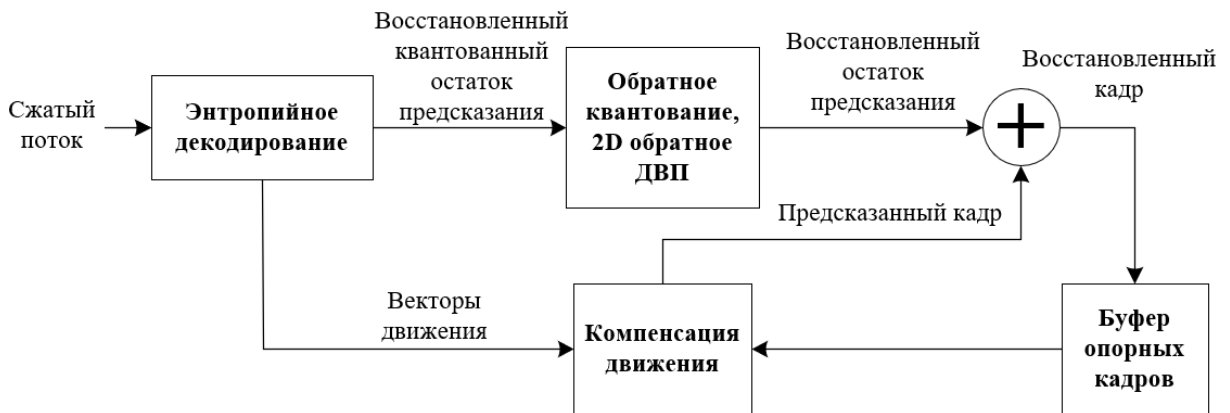


Рисунок 2. Общая схема вейвлет-видеодекодера

Во второй главе рассмотрена возможность применения комплексного вейвлет-преобразования (КВП) для задачи компенсации движения. Для сравнения эффективности предложенного метода с использованием КВП в этой главе приведены основные принципы работы самых распространённых методов компенсации движения – сопоставления блоков (СБ) и сопоставления перекрывающихся блоков (СПБ).

Затем в данной главе обоснованы актуальность и новизна КВП для задачи сжатия подвижных изображения. В связи с трудностью построения обратимого КВП и вычисления с комплексными числами, КВП удобно построить на основе двойного дерева комплексного вейвлет-преобразования (ДДКВП), которое в одномерном случае состоит из двух действительных ортогональных или биортогональных вейвлет-базисов, один из которых отвечает за «действительную часть», а второй за «мнимую часть», то есть, с одной стороны, используются действительные вейвлет-коэффициенты, а с другой – вейвлет-преобразование носит комплексный характер.

Далее приведено построение одномерного и двумерного ДДКВП с помощью банка фильтров вейвлет-преобразования. Теоретически и практически получен метод вычисления комплексных коэффициентов ДДКВП, из которых также можно определить амплитуды и фазы коэффициентов ДДКВП, использованные в предложенном методе компенсации движения.

После этого в данной главе предложен метод компенсации движения, а именно, метод построения предсказанного кадра, использующий ДДКВП, который представлен на рисунке 3.

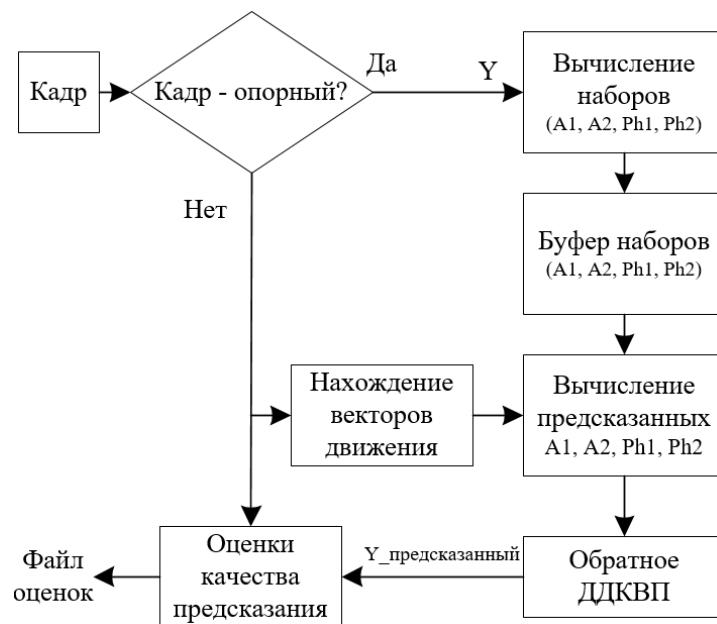


Рисунок 3. Алгоритм построения предсказанного кадра

На вход последовательно поступают кадры. Если кадр опорный, то сохраняются амплитуды и фазы всех возможных смещенных кадров с точностью до четверти пиксела, иначе выполняется процесс нахождения векторов движения с точностью от одного до четверти пиксела методом СБ с помощью алгоритма полного перебора. Затем предложенный метод определяет амплитуды и фазы каждого блока предсказанного кадра на основе найденных векторов движения и буфера наборов амплитуд и фаз комплексных коэффициентов. После чего вычисляется предсказанный кадр обратным ДДКВП. Блок «оценки качества предсказания» получает предсказанный кадр, вычисляет метрики для сравнения и записывает результаты в файл оценок.

Для проверки работоспособности и оценки эффективности предложенного метода была создана экспериментальная программа, написанная на C++, выполняющая чтение видео, поиск векторов движения с точностью до четверти пиксела методом полного перебора, построение предсказанных кадров разными методами с помощью опорного кадра и найденным набором векторов движения, запись предсказанного кадра и остатков предсказания для визуального сравнения методов предсказания, и сравнение эффективности методов разными метриками. Нужно отметить, что экспериментальная программа обрабатывает данные как вейвлет-видеокодер, схема которого представлена на рисунке 1.

В ходе работы с видео высокого разрешения, в частности, с видеопоследовательностями «720p_city» и «720p_stockholm», были получены следующие основные результаты:

- Качество предсказания предложенного метода мало зависит от типа вейвлета, что позволяет построить ДДКВП с любым вейвлет-фильтром и легко адаптировать предложенный метод в существующих вейвлет-видеокодеках.

- Предложенный метод обеспечивает реализацию меньшей энергии остатка предсказания. Конкретно, если отклонение между текущим и опорным кадрами мало, то предложенный метод выигрывает по сравнению с эталонными методами от 0.2 до 0.5 дБ PSNR остатка предсказания, иначе предложенный метод выигрывает больше 0.5 дБ PSNR остатка предсказания.

- При одинаковом качестве восстановленного кадра предложенный метод снижает объем сжатого потока для квантованного остатка предсказания от 20 до 25% для видео «720p_city» и от 6 до 8% для видео «720p_stockholm». Следовательно, предложенный метод лучше эталонных методов подходит для сжатия информации.

В третьей главе приведен принцип работы традиционного метода компенсации движения, основанного на векторах движения, найденных с функцией стоимости, вычисленной по формуле:

$$SAD(V_x, V_y) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left| Y(x+m, y+n) - X(x+V_x+m, y+V_y+n) \right|. \quad (1)$$

Здесь $Y(u, v)$ - значение яркости пикселя с координатами (u, v) в текущем кадре;

$X(u, v)$ - значение яркости пикселя с координатами (u, v) в опорном кадре;

$M \times N$ - размер блока;

(V_x, V_y) - вектор движения;

(x, y) – координаты верхнего левого угла рассматриваемого блока в текущем кадре.

Приведенная выше функция легко вычисляется, однако она соответствует только плоскопараллельному перемещению при анализе и компенсации движения. В настоящей главе предложена новая функция стоимости, вычисляемая по формуле:

$$SAD_F(V_x, V_y) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left| Y(x+m, y+n) - F\left(X(x+V_x+m, y+V_y+n)\right) \right|. \quad (2)$$

В формуле (2) введена функция $F(X(x, y))$, называемая функцией аппроксимации. В работе исследованы 8 элементарных функций аппроксимации, представленных в таблице 1. По введенной терминологии традиционный метод компенсации движения использует простую функцию аппроксимации $F(X(x, y)) = X(x, y)$, представленную в таблице 1 под номером 0, а предложенный алгоритм применяет более сложные функции.

Таблица 1. Список исследованных функций аппроксимации

Номер функции	Функции аппроксимации
0	$F(X) = X$
1	$F(X) = aX + b$
2	$F(X) = bX^a$
3	$F(X) = be^{aX}$
4	$F(X) = a \cdot \ln(X) + b$
5	$F(X) = (a/X) + b$
6	$F(X) = X/(bX + a)$
7	$F(X) = bX^{aX}$
8	$F(X) = bX^{a/X}$

Для применения предложенного метода в этой главе были решены и подробно описаны задачи, связанные с дополнительными параметрами введенных функций аппроксимации, такие как метод определения, требуемая точность и способ передачи дополнительных параметров. Также для всех функций аппроксимации оценена точность представления дополнительных параметров, обеспечивающая достаточно малую ошибку предсказания значения яркости текущего кадра из-за квантования дополнительных параметров. Как и вектор движения, каждая компонента дополнительных параметров передается отдельно в связи с тем, что она принимает значения из определенного диапазона и эти значения в какой-то степени оказываются коррелированы. Для разработчиков видеокодеков также приведены некоторые

алгоритмы оптимизации предложенного метода, связанные с требованием снижения вычислительной сложности видеокodeка.

С целью проверки работоспособности и эффективности предложенного метода разработана экспериментальная программа, написанная на языке программирования C++, которая оценивает качество восстановления и количество битов, необходимых для кодирования видео по традиционному методу анализа и компенсации движения и по предложенному способу. Проведенный анализ результатов обработки сложных видеопоследовательностей показал, что исследованные методы выигрывают по качеству у традиционного метода и среди исследованных функций аппроксимации оптимальной является линейная функция аппроксимации, представленная в таблице 1 под номером 1.

После подтверждения эффективности предложенного способа с помощью экспериментальной программы для проверки практической реализации предложенный способ был внедрен в вейвлет-видеокodeк Divac, обзор которого приведен в главе 1. После адаптации предложенного способа в рассматриваемом видеокodeке было проведено тестирование на ряде видео различных типов для разных режимов разбиения кадра на блоки размером 32×32 , 16×16 и 8×8 пикселей. Результаты анализа видеокodeка Divac также показали, что лучше всего использовать линейную функцию аппроксимации в предложенном методе. Далее приведены результаты применения предложенного метода с использованием линейной функции аппроксимации. В качестве примера на рисунках 4 и 5 представлены кривые зависимости уровня искажений от скорости выходного потока (RD-кривые) для видео «1080p_riverbed» в режимах разбиения кадра на блоки 16×16 и 32×32 .

Проведенные численные сравнения скорости выходного потока при заданном качестве восстановления для ряда видео различных типов в разных случаях размера блока представлены в таблице 2.

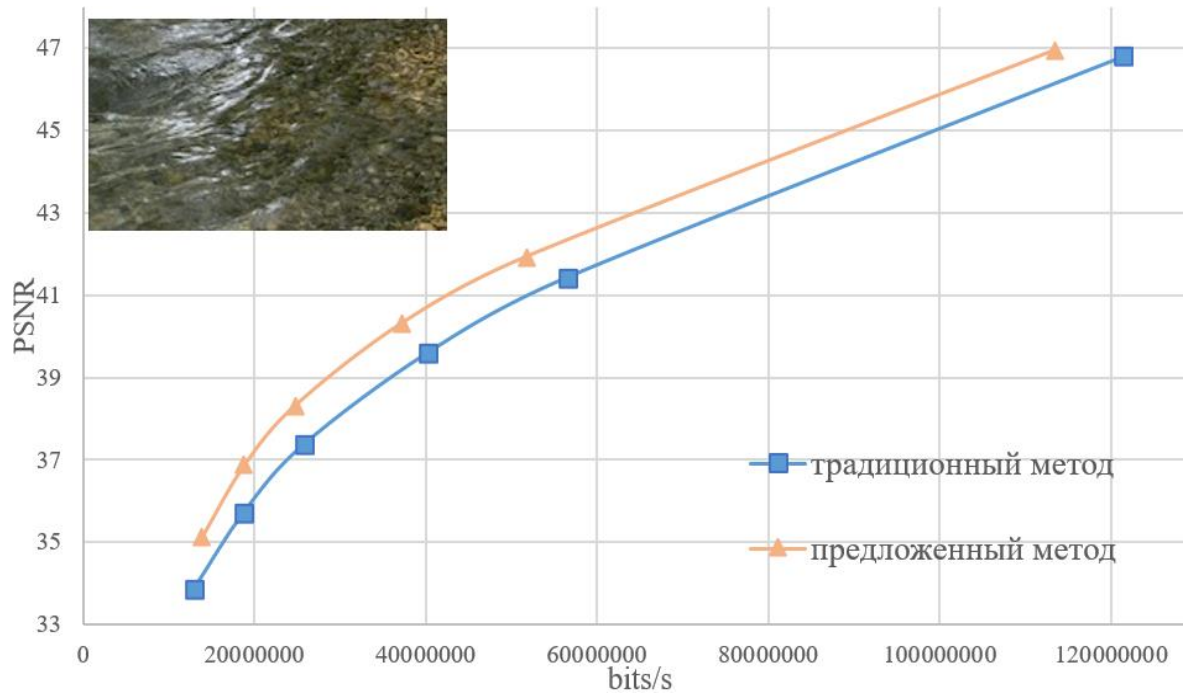


Рисунок 4. Сравнение RD-кривых предложенного и традиционного методов для видео «1080p_riverbed» в случае размера блока 16×16

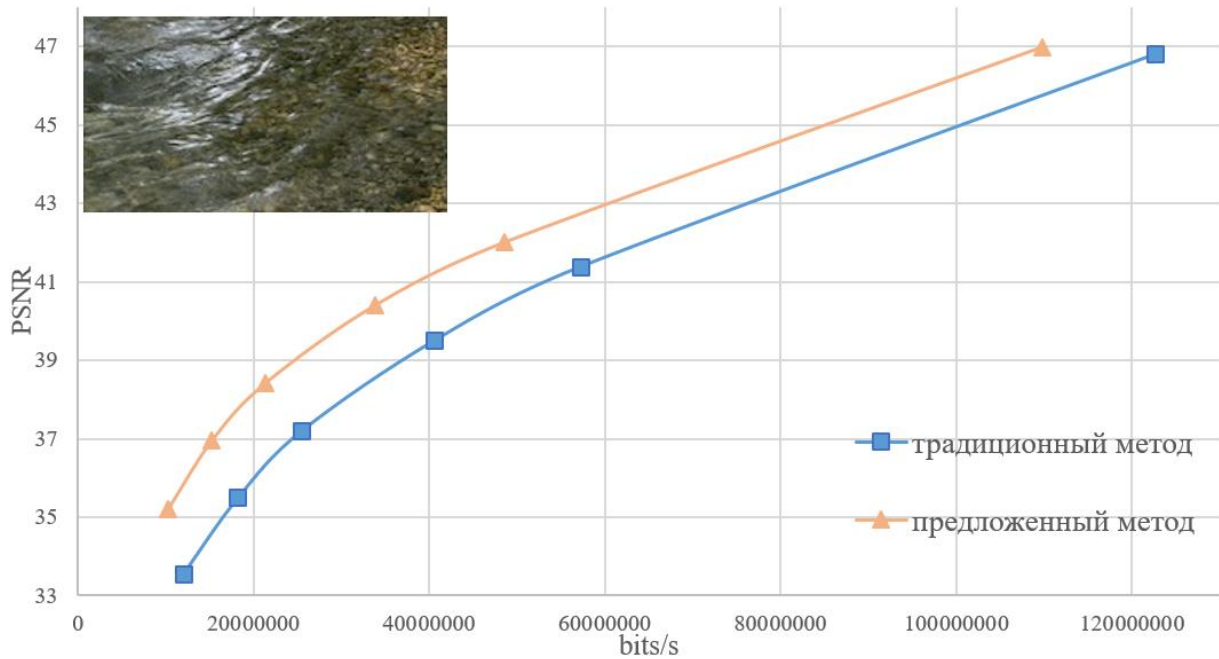


Рисунок 5. Сравнение RD-кривых предложенного и традиционного методов для видео «1080p_riverbed» в случае размера блока 32×32

Таблица 2. Выигрыш предложенного способа в сравнении с традиционным по скорости выходного потока при заданном качестве восстановления

Размер блока	16×16		32×32	
PSNR, дБ	40	42	40	42
видео				
1080p_blue_sky	-	2%	5.5%	9%
1080p_pedestrian_area	10%	20%	33%	29%
1080p_riverbed	19.5%	17%	30%	25%
1080p_rush_hour	-	33%	20%	40%
1080p_station2	0	12.5%	19%	15%
1080p_sunflower	-	9%	12.5%	12%
1080p_tractor	3%	7.5%	17%	14%
720p_mobcal_ter	2%	2%	4.5%	2.5%
720p_shields_ter	2%	4%	3%	2%
720p_stockholm	3%	2%	4%	3%
4cif_crew	17%	14%	22%	17%
4cif_soccer	4%	4%	8.5%	7%

Из таблицы 2 и из построенных графиков можно сделать следующие выводы:

- Среди исследованных функций аппроксимации наилучшей является линейная функция аппроксимации, представленная в таблице 1 под номером 1. В общем случае предложенный метод с различными функциями аппроксимации выигрывает у традиционного метода при хорошем качестве восстановления (например, PSNR=40 дБ) и при высоком качестве восстановления (например, PSNR=42 дБ).
- В общем случае RD-кривые предложенного и эталонного методов пересекаются в точке со значением PSNR где-то от 35 дБ до 37 дБ, то есть, для участка RD-кривой с низким качеством предложенный способ проигрывает

традиционному, однако этот участок не является интересующим для многих приложений сжатия видео.

- Для видео высокого разрешения 1080p (1920x1080) предложенный способ существенно уменьшает выходной поток при заданном качестве восстановления, что также объясняет актуальность и практическую значимость предложенного способа в связи с тем, что требования пользователей увеличиваются и часто нужно передавать видео высокого разрешения с требованием высокого качества восстановления.

- Чем сложнее тестовые видео для предсказания (изменения в сцене отличаются от плоскопараллельного перемещения), тем лучше предложенный способ по сравнению с традиционным.

- Для случая размера блока 16×16 , предложенный способ может сократить выходной поток до 20% при хорошем качестве восстановления PSNR = 40 дБ и до 33% при отличном качестве PSNR = 42 дБ. А для случая размера блока 32×32 применение предложенного метода позволит уменьшить выходной поток от 3% до 33% для качества PSNR = 40 дБ и от 2% до 40% для качества PSNR = 42 дБ.

В конце главы приведены некоторые подходы к оптимизации предложенного способа, направленные на уменьшение количества данных, приходящихся на кодирование дополнительных параметров. Проведенным исследованием получены оптимальная точность и наиболее подходящий метод их кодирования для каждого режима разбиения кадра на блоки. Из проведенного анализа для предложенного способа после оптимизации можно сделать следующие выводы:

- В сравнении с предложенным способом до оптимизации при заданном качестве восстановления оптимизация предложенного способа может существенно сократить объем данных (от 4% до 32% для качества PSNR = 40 дБ, от 3% до 26% для качества PSNR = 42 дБ) для режима разбиения кадра на

малые блоки 8×8 , а для других размеров блока, оптимизация обеспечивает небольшое уменьшение выходного потока (до 10% при PSNR = 40 дБ, до 6.5% при PSNR = 42 дБ для размера блока 16×16 , до 3% при PSNR = 40 дБ, до 1.3% при PSNR = 42 дБ для размера блока 32×32).

- В сравнении с традиционным методом, предложенный вариант после оптимизации может сократить выходной поток до 20% при хорошем качестве восстановления PSNR = 40 дБ и до 33.5% при отличном качестве PSNR = 42 дБ для случая размера блока 16×16 . А в случае размера блок 32×32 , применение предложенного способа позволит уменьшить выходной поток от 3% до 35% для качества PSNR = 40 дБ и от 2% до 40% для качества PSNR = 42 дБ. Нужно также отметить, что по сравнению с аналогичным выводом для начальной версии предложенного метода до оптимизации, здесь увеличивается до 5% процент выигрыша для каждого тестового видео.

В четвертой главе обоснованы актуальность и новизна исследования метода квантования для вейвлет-видеокодека Dirac и поставлена цель этой части диссертационной работы – исследование и разработка оптимизированного метода квантования в блоках «квантование» и «обратное квантование» для модифицированного видеокодека Dirac, схемы которого показана на рисунках 1 и 2.

В модифицированной версии видеокодека Dirac использовано многоканальное вейвлет-разложение в блоке «двумерное дискретное вейвлет-преобразование», представленном на рисунке 1. Один уровень трехканального ДВП позволяет разложить одномерный сигнал на три диапазона, а двумерный сигнал – на девять диапазонов. На рисунке 6 показано разделение двумерного сигнала на диапазоны для одного уровня трехканального вейвлет-разложения.

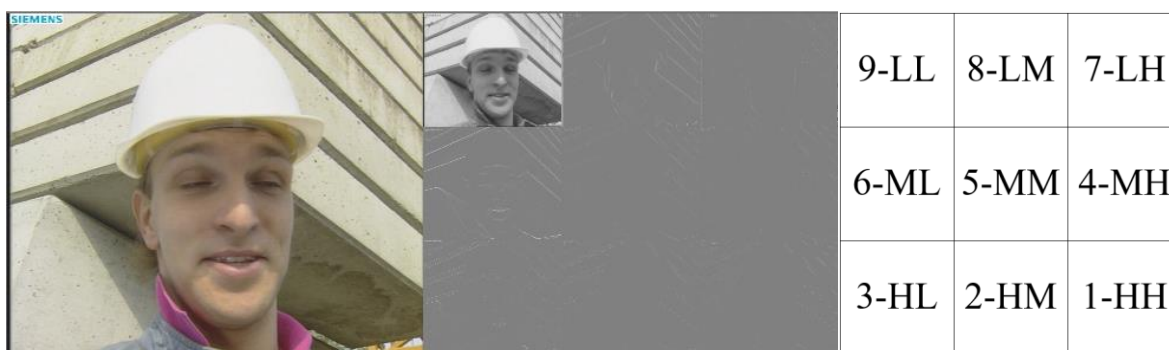


Рисунок 6. Разделение сигнала на диапазоны с помощью трехканального вейвлет-разложения для первого уровня

Особенность многоканального вейвлет-разложения заключается в том, что после фильтрации большая часть энергии сигнала сосредоточивается в низкочастотном диапазоне по горизонтали и по вертикали (9-LL), а остальная энергия сигнала распределяется между остальными диапазонами неравномерно, из чего следует различное влияние искажений вейвлет-коэффициентов разных диапазонов на восстановленный сигнал, что позволяет квантовать диапазоны по-разному.

В данной части диссертационной работы исследованы особенности многоканального вейвлет-разложения и различные алгоритмы квантования, такие как алгоритм неравномерного квантования Ллойда-Макса, алгоритм энтропийного ограниченного квантования и их потенциальные возможности применения для задачи сжатия видео. На основе результатов этих исследований для вейвлет-видеокодека Dirac предложен оптимизированный метод квантования, основанный на совместном применении известного метода равномерного квантования для высокочастотных диапазонов (1-NH, 2-NM, 4-MH) (см. рисунок 6) и алгоритма Ллойда-Макса для других диапазонов.

В качестве примера на рисунках 7 и 8 представлена оценка предложенного метода по сравнению с эталонным методом равномерного квантования, использованным в видеокодеке Dirac до внедрения предложенного метода.

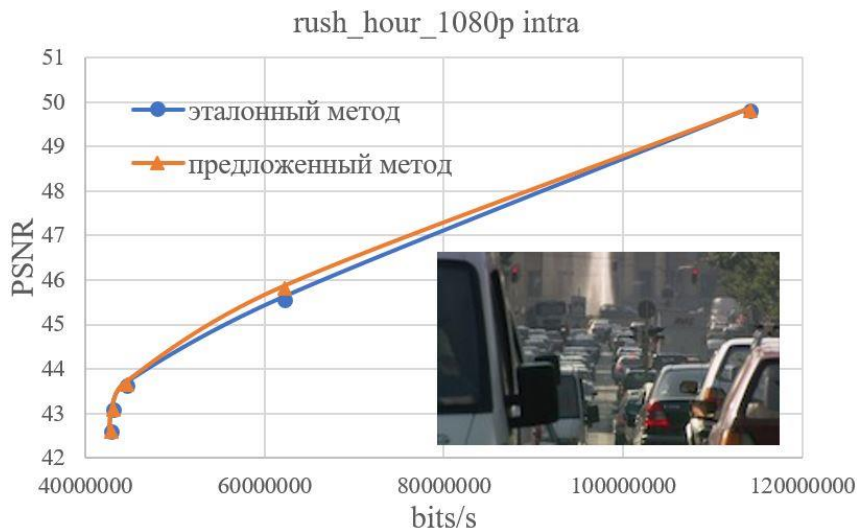


Рисунок 7. Оценка предложенного метода для видео «rush_hour_1080p» в режиме INTRA

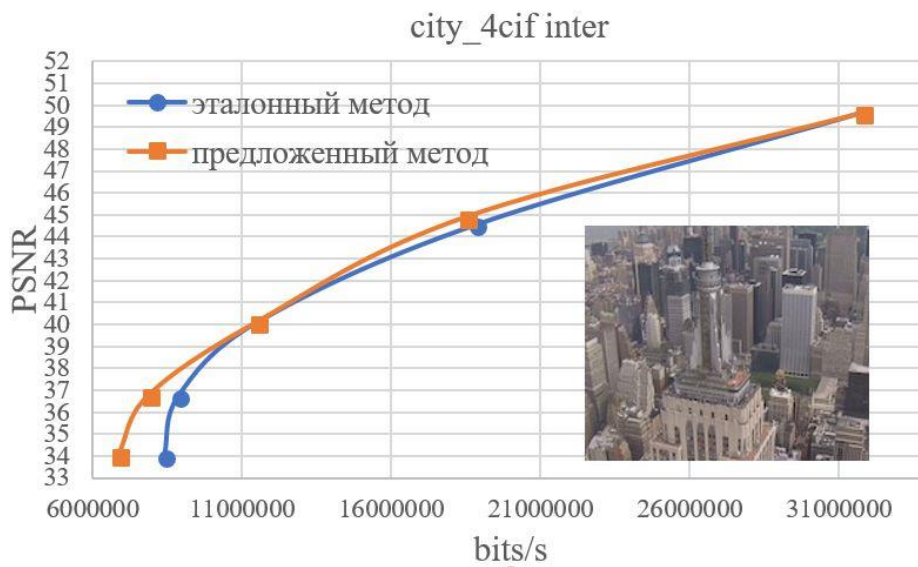


Рисунок 8. Оценка предложенного метода для видео «city_4cif» в режиме INTER

В ходе работы с видео разных типов получены следующие выводы:

- Если приходится использовать линейное квантование для группы каких-либо диапазонов, то оптимальным вариантом для этих диапазонов является использование равномерного квантования, т.е. использование линейного

квантования с одинаковым шагом квантования для рассматриваемых диапазонов.

- Алгоритм энтропийного ограниченного квантования дает результат, примерно равный алгоритму Ллойда-Макса для первой итерации, так что достаточно и целесообразно остановиться на использовании алгоритма Ллойда-Макса.

- Предложен метод оптимизированного квантования на основе совместного применения известного метода равномерного квантования для высокочастотных диапазонов (1-НН, 2-НМ, 4-МН) (см. рисунок 6) и алгоритма Ллойда-Макса для других диапазонов.

- Для режима внутрикадрового кодирования предложенный метод дает примерно равный результат на областях низкого качества RD кривой и немного лучше для областей высокого качества RD кривой по сравнению с методом равномерного квантования, реализованного в рассматриваемом видеокодеке Divac, конкретно, предложенный метод уменьшает до 4.2% битовой скорости.

- Для режима межкадрового кодирования предложенный метод всегда дает значительно лучший результат по сравнению с равномерным квантованием, а именно, предложенный метод уменьшает до 11.2% битовой скорости на участке с приемлемым качеством ($PSNR = 37$ дБ) RD кривой, а на участке высокого качества RD кривой такой выигрыш составляет до 9.8%.

- Более того, для сложных видео, когда межкадровое предсказание работает не очень хорошо, предложенный метод работает намного лучше эталонного в областях с приемлемым PSNR, что означает целесообразность применения предложенного метода по сравнению с эталонным.

ЗАКЛЮЧЕНИЕ

В настоящей диссертационной работе исследованы и разработаны методы сжатия подвижных изображений с помощью многоканального и комплексного вейвлет-разложений. Для усовершенствования методов сжатия подвижных изображений в общем и развития вейвлет-видеокодека Dirac в частности в работе поставлены и выполнены некоторые исследования, направленные на актуальные задачи – поиска и компенсации движения и системы квантования. В ходе исследований получены следующие основные научные и практические результаты:

1. Предложен метод компенсации движения на основе комплексного вейвлет-разложения. По оценке энтропии выходного потока для видео высокого разрешения 720p (1280x720) в разработанной экспериментальной программе, написанной на языке программирования C++, получено, что при одинаковом качестве восстановленного кадра предложенный метод снижает объем сжатого потока для квантованного остатка предсказания от 6% до 25% в сравнении с эталонными методами сопоставления блоков и сопоставления перекрывающихся блоков, которые широко применяются в современных видеокодеках. Кроме того, любой вейвлет-фильтр может быть применен для построения комплексного вейвлет-разложения, примененного в предложенном методе, так что предложенный метод может быть применен в существующих вейвлет-видеокодеках.

2. Предложен способ улучшения качества традиционного метода компенсации движения с помощью новой функции стоимости, в которой введена функция с дополнительными параметрами, называемая функцией аппроксимации. Для разработки предложенного способа исследованы 8 элементарных функций аппроксимации и теоретически решены задачи, связанные с дополнительными параметрами введенных функций аппроксимации, такие как метод определения, требуемая точность и способ передачи дополнительных параметров.

3. Разработана экспериментальная программа, написанная на языке программирования C++, оценивающая качество предложенного способа улучшения качества традиционного метода компенсации движения. По оценке энтропии выходного потока, в сравнении с традиционным предложенный способ существенно уменьшает объем выходного потока при заданном качестве восстановления для видео высокого разрешения 1080p (1920x1080), при этом линейная функция аппроксимации является оптимальной среди исследованных функций.

4. Предложенный способ улучшения качества традиционного метода компенсации движения внедрен в вейвлет-видеокодек Dirac. Применение этого способа в рассматриваемом видеокодеке может уменьшить выходной поток до 20% при хорошем качестве восстановления $PSNR = 40$ дБ и до 33% при отличном качестве $PSNR = 42$ дБ для случая разбиения кадра на блоки 16×16 . А для случая размера блока 32×32 , применение предложенного метода позволит уменьшить выходной поток от 3% до 33% для качества $PSNR = 40$ дБ и от 2% до 40% для качества $PSNR = 42$ дБ. Кроме того, в сравнении с традиционным предложенный способ значительно удачнее работает при сложных видеопоследовательностях, в которых происходят изменение масштаба (зуммирование), вращение, неравномерное изменение яркости и т.п.

5. Предложены и разработаны некоторые подходы к оптимизации способа улучшения качества традиционного метода компенсации движения как на уровне качества способа, так и на уровне его реализации. Для этого исследованы некоторые подходы, связанные с уменьшением точности передачи дополнительных параметров и самым подходящим методом их кодирования. Проведенным анализом получены оптимальные точности параметров для разных режимов разбиения кадра на блоки и лучший метод их кодирования с пространственным предсказанием.

6. Исследованные алгоритмы оптимизации предложенного способа повышения качества традиционного метода компенсации движения позволяют существенно сократить объем данных в сравнении с предложенным способом

до оптимизации, конкретно от 4% до 32% при качестве восстановления PSNR = 40 дБ, от 3% до 26% при качестве PSNR = 42 дБ для режима разбиения кадра на малые блоки 8×8, до 10% при PSNR = 40 дБ, до 6.5% при PSNR = 42 дБ для размера блока 16×16, до 3% при PSNR = 40 дБ, до 1.3% при PSNR = 42 дБ для размера блока 32×32. По сравнению с традиционным методом, предложенный способ после оптимизации может сократить выходной поток до 20% при хорошем качестве PSNR = 40 дБ и до 33.5% при отличном качестве PSNR = 42 дБ для случая размера блока 16×16. А в случае размера блока 32×32, применение предложенного метода позволит уменьшить выходной поток от 3% до 35% для качества PSNR = 40 дБ и от 2% до 40% для качества PSNR = 42 дБ. Нужно также отметить, что по сравнению с результатом 4, здесь меняются не только пределы достижения предложенного способа, но и увеличивается процент выигрыша до 5% для каждого тестового видео.

7. Исследованы особенности многоканального вейвлет-разложения и различные алгоритмы квантования, такие как алгоритм неравномерного квантования Ллойда-Макса, алгоритм энтропийного ограниченного квантования и их потенциальные возможности применения для задачи сжатия видео. На основе результатов этих исследований предложен оптимизированный метод квантования для вейвлет-видеокодека Dirac, который основан на совместном применении известного метода равномерного квантования для некоторых высокочастотных диапазонов и алгоритма Ллойда-Макса для других диапазонов.

8. Предложенный метод квантования внедрен в вейвлет-видеокодеке Dirac. Для режима внутрикадрового кодирования предложенный метод дает примерно равный результат в областях низкого качества RD кривой и немного лучше для областей высокого качества RD кривой по сравнению с методом равномерного квантования, реализованного в рассматриваемом видеокодеке Dirac. Предложенный метод уменьшает до 4.2% битовой скорости. Для режима межкадрового кодирования, предложенный метод всегда дает значительно лучший результат по сравнению с равномерным квантованием. Предложенный

метод уменьшает до 11.2% битовой скорости на участке с приемлемым качеством (PSNR = 37 дБ) RD кривой, а на участке высокого качества RD кривой такой процент составляет до 9.8%. Более того, для сложных видео, когда межкадровое предсказание работает не очень хорошо, предложенный метод намного лучше эталонного в областях с приемлемым PSNR, что означает целесообразность использования предложенного метода квантования по сравнению с эталонным, реализованным в рассматриваемом видеокодеке до адаптации предложенного метода.

Основные публикации по теме диссертации

Публикации в изданиях, входящих в перечень RSCI или индексируемых в международных базах данных (SCOPUS, Web Of Science)

1. Дам Чонг Нам. Оптимизация метода квантования для вейвлет-видеокодека // Цифровая обработка сигналов. 2018. №3. С. 44-48.
2. Дам Чонг Нам. Применения комплексного вейвлет-преобразования в задаче сжатия видео // Цифровая обработка сигналов. 2018. №3. С. 38-43.

Публикации, принятые к печати в изданиях, входящих в перечень RSCI или индексируемых в международных базах данных (SCOPUS, Web Of Science)

1. Dam Trong Nam, Gennady Yu. Gryzov, Alexander V. Dvorkovich, Viktor P. Dvorkovich. Nonlinear Quantization Method for Wavelet-Based Video Codec // 5th International Conference «Engineering and Telecommunication – En&T-2018». Moscow, Russia, Nov. 2018.
2. Дам Чонг Нам. Особенности реализации модифицированного метода компенсации движения // Цифровая обработка сигналов. 2019. №3.
3. Дам Чонг Нам. Улучшение качества традиционного метода компенсации движения // Цифровая обработка сигналов. 2019. №2.

Заявка на патент

1. Дворкович А.В., Грызов Г.Ю., Дам Чонг Нам. Заявка на изобретение. Способ компенсации движения и устройство для его реализации // Роспатент, регистрационный №2019111048.

Публикации в материалах конференций

1. Дам Чонг Нам. Компенсация движения с помощью комплексного вейвлет-преобразования в задаче сжатия видео // 20-я международная конференция «Цифровая обработка сигналов и ее применение – DSPA 2018», Москва, Россия, доклады. Том 2. ст. 578-583.

2. Дам Чонг Нам. Предсказание кадра с помощью комплексного вейвлет-преобразования в задаче сжатия видео // 20-я международная конференция «Цифровая обработка сигналов и ее применение – DSPA 2018», Москва, Россия, доклады. Том 2. ст. 573-578.

3. Дам Чонг Нам. Способ повышения качества в традиционном методе компенсации движения // Труды 61-й Всероссийской научной конференции МФТИ. 19–25 ноября 2018 года. Радиотехника и компьютерные технологии. – М.: МФТИ, 2018. -168 с. ст. 129-130.