

На правах рукописи



Веселов Антон Игоревич

**ОБРАБОТКА ВИДЕОИНФОРМАЦИИ В СИСТЕМАХ СЖАТИЯ,
ОСНОВАННЫХ НА ПРИНЦИПАХ КОДИРОВАНИЯ ЗАВИСИМЫХ
ИСТОЧНИКОВ**

Специальность 05.12.13 — «Системы, сети и устройства телекоммуникаций»

Автореферат
диссертации на соискание учёной степени
кандидата технических наук

Санкт-Петербург — 2015

Работа выполнена в Федеральном государственном автономном образовательном учреждении высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения»

Научный руководитель: доктор технических наук, доцент
Тюрликов Андрей Михайлович

Официальные оппоненты: **Кудряшов Борис Давидович**
доктор технических наук, профессор кафедры «Информационные системы» Федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский национальный исследовательский университет информационных технологий»

Зубакин Игорь Александрович
кандидат технических наук, доцент кафедры «Телевидение и видеотехника» Федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)»

Ведущая организация: **Открытое акционерное общество «Научно-исследовательский институт телевидения» (ОАО «НИИТ»)**, г. Санкт-Петербург

Защита состоится 15 марта 2016 г. в 14-00 на заседании диссертационного совета Д 212.233.05 в Федеральном государственном автономном образовательном учреждении высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения» по адресу: г. Санкт-Петербург, ул. Б. Морская, 67, ауд. 53-01.

С диссертацией можно ознакомиться в библиотеке ФГАОУ ВО ГУАП и на сайте www.guar.ru.

Автореферат разослан 15 января 2016 г.

Ученый секретарь
диссертационного совета



Овчинников Андрей Анатольевич
к.т.н., доцент

Общая характеристика работы

Актуальность темы. В последнее время все большее распространение получает передача видеоданных от мобильных устройств по беспроводным сетям связи. Это связано как с существенным развитием технологий беспроводной передачи данных, так и с развитием мобильных устройств: беспроводных сенсоров, беспроводных камер видеонаблюдения, мобильных пользовательских устройств и т. д. В таких системах передатчик, как правило, характеризуется малыми вычислительными мощностями и ограниченной емкостью аккумуляторного устройства. При этом существенные ограничения накладываются как на мощность процессора, так и на объем памяти.

Существующие технологии сжатия видеоданных, в первую очередь подходы, описанные в стандартах серий ITU-T H.26x и ISO/IEC MPEG, не учитывают специфику источника информации в подобных системах, поэтому разработка новых методов сжатия визуальных данных, не требующих больших вычислительных затрат на стороне передатчика, является важной и актуальной задачей.

В диссертационной работе рассматривается распределенное кодирование видеоданных – подход к сжатию, основанный на методах кодирования зависимых источников, позволяющий существенно снизить сложность обработки на стороне передатчика. Решается **задача** эффективного восстановления промежуточных кадров на стороне декодера.

Различные аспекты кодирования зависимых источников (или распределенного кодирования источников) представлены в работах известных отечественных и зарубежных авторов (С.И.Гельфанд, В.Д.Колесник, Б.Д.Кудряшов, М.С.Пинскер, Г.Ш.Полтырев, А.Вайнер, Д.Вулф, Д.Слепян). Однако, до недавнего времени практическая реализация этих идей не была востребована. Только в конце 1990-х годов появились прикладные задачи, в которых применение распределенного кодирования могло дать преимущества по сравнению с существующими на тот момент подходами. В качестве перспективной области рассматривалось сжатие видеоданных в системах с ограниченными вычислительными ресурсами на стороне передатчика информации. Одним из существенных преимуществ от применения распределенного кодирования в данной системе, является то, что процедуру оценки и устранения временной избыточности, основанную на предсказании промежуточных кадров, можно выполнять

на приемнике, за счет чего существенно снижается сложность обработки на передатчике. Процедуру межкадрового предсказания в таких системах принято называть *генерацией дополнительной информации*. Кроме того, т. к. задача восстановления промежуточных кадров решается декодером, повышение степени сжатия возможно только за счет модификации приемника. В рамках данного класса прикладных задач были разработаны концепции кодеков, основанные на принципах распределенного кодирования источников. В последнее время большое внимание уделяется расширению и модификации этих концепций с учетом появляющихся новых прикладных задач, таких как эффективное кодирование многомерных и многокадровых видеопоследовательностей. Несмотря на это, в базовых принципах остается ряд открытых вопросов. К их числу следует отнести учет особенностей входных данных для процедуры межкадрового предсказания на стороне декодера, а также оценку параметров ошибок межкадрового предсказания. Также следует отметить, что в большинстве современных работ приводится описание эвристических подходов для решения практических задач, возникающих при распределенном кодировании визуальных данных. При этом теоретические модели данных исследованы не в полной мере, что не позволяет эффективно разрабатывать новые и улучшать существующие алгоритмы сжатия, использующие данный подход.

Целью диссертационного исследования является повышение степени сжатия без ухудшения качества восстановления видеоданных в кодеках, основанных на принципах кодирования зависимых источников с дополнительной информацией на декодере, за счет усовершенствования существующих и разработки новых способов обработки данных на стороне декодера.

Для достижения поставленной цели необходимо решить следующие **задачи**:

1. Исследовать типовые методы сжатия видеоданных, основанные на принципах распределенного кодирования источников.
2. Исследовать способы формирования дополнительной информации на стороне декодера.
3. Предложить новый алгоритм генерации дополнительной информации, учитывающий особенности входных данных.

4. Исследовать статистические характеристики ошибок межкадрового предсказания, возникающих при генерации дополнительной информации.
5. Разработать модель ошибок предсказания промежуточных кадров на стороне декодера.
6. Предложить алгоритм оценки параметров ошибок межкадрового предсказания.

Объект и предмет исследования. Объектом исследования является система сжатия видеоданных, основанная на принципах кодирования зависимых источников с дополнительной информацией на декодере.

Предмет исследования составляет процесс восстановления промежуточных кадров на стороне декодера.

Методы исследования. При получении основных результатов работы использовались общие методы системного анализа, методы теории вероятностей и математической статистики, теории случайных процессов, в частности марковских случайных полей, теории принятия решений, методы машинного зрения, а также методы имитационного моделирования.

Научная новизна:

1. Построена модель системы сжатия на базе распределенного кодирования для анализа алгоритмов восстановления промежуточных кадров на стороне декодера, позволяющая в одинаковых условиях производить сравнение различных методов сжатия видеоданных.
2. Предложен алгоритм генерации дополнительной информации декодера, отличающийся от существующих алгоритмов тем, что использует метод оценки движения, основанный на модели истинного движения объектов в видеоданных.
3. Предложен способ сравнительной оценки различных алгоритмов межкадрового предсказания в системах распределенного кодирования видеоданных, отличающийся от существующих тем, что позволяет в одинаковых условиях производить сравнительную оценку методов предсказания, не учитывая влияние прочих методов системы сжатия.

4. Впервые предложена модель ошибок предсказания промежуточных кадров, в явном виде учитывающая пространственные особенности артефактов интерполяции.
5. Предложен алгоритм оценки параметров ошибок предсказания промежуточных кадров, отличающийся от существующих тем, что позволяет в явном виде учитывать пространственную зависимость между ошибками.

Практическая значимость диссертационной работы. Полученные в диссертационной работе результаты позволяют повысить по сравнению с существующими алгоритмами эффективность сжатия видеоданных в системах, основанных на распределенном кодировании источников, что способствует уменьшению энергопотребления и габаритных размеров кодеров видеoinформации.

Достоверность. Результаты, полученные в диссертационной работе, согласуются с известными теоретическими моделями распределенного кодирования. Основные результаты опубликованы в рецензируемых журналах и доложены на крупных международных конференциях.

Апробация работы. Основные результаты работы докладывались и обсуждались на следующих конференциях и симпозиумах в период с 2009 по 2014 гг.: на научных сессиях ГУАП; на 14-й конференции «IEEE Multimedia Signal Processing»; на 6-й конференции «International Congress on Ultra Modern Telecommunications and Control Systems»; на 13-м симпозиуме «Problems of Redundancy in Information and Control Systems»; на 16-й конференции «IEEE Multimedia Signal Processing»; на 8-й международной конференции «KES Conference on Intelligent Interactive Multimedia: Systems And Services».

Внедрение результатов. Результаты работы были использованы в рамках проекта «Разработка цепочки фильтров постобработки видеоданных», осуществляемого ЗАО «Интел А/О». Кроме того, результаты работы используются в учебном процессе кафедры инфокоммуникационных систем ГУАП.

Личный вклад. Все результаты, представленные в тексте диссертационной работы, получены автором лично.

Публикации. Материалы, отражающие основное содержание и результаты диссертационной работы, опубликованы в 15 печатных работах. Из них

2 работы опубликованы в рецензируемых научных журналах, утвержденных в перечне ВАК, и 5 работ опубликованы в журналах, индексируемых в Scopus.

Основные положения, выносимые на защиту:

1. Алгоритм межкадрового предсказания для кодеков видеоинформации, основанных на принципах кодирования зависимых источников с дополнительной информацией на декодере, позволяющий уменьшить по сравнению с существующими алгоритмами число ошибок предсказания за счет использования временной интерполяции с учетом истинного движения объектов.
2. Модель кодека без обратной связи, позволяющая производить сравнение алгоритмов межкадрового предсказания в системах кодирования зависимых источников видеоинформации с дополнительной информацией на декодере.
3. Модифицированный алгоритм оценки параметров ошибок межкадрового предсказания в спектральной области, который за счет учета неоднородности ошибок в спектральных коэффициентах позволяет уменьшить битовые затраты на восстановление промежуточного кадра.
4. Модель ошибок межкадрового предсказания, основанная на Марковских случайных полях, которая позволяет учитывать пространственную зависимость между ошибками в спектральных коэффициентах.

Объем и структура работы. Диссертация состоит из введения, четырех разделов, заключения и двух приложений. Полный объем диссертации составляет 142 страницы с 28 рисунками и 7 таблицами. Список литературы содержит 66 наименований.

Содержание работы

Во *введении* обоснована актуальность исследования методов понижения сложности кодирования видеоданных на передающей стороне, представлена научная новизна диссертационной работы и ее практическая ценность сформулированы основные положения, выносимые на защиту, а также приведено краткое содержание диссертационной работы по разделам.

В *первом разделе* дается краткое введение в теорию распределенного кодирования источников, указываются базовые теоретические результаты, а также даются основные определения и понятия, используемые в работе.

В работе рассматривается процесс обработки видеоданных в системе сжатия, использующей принципы кодирования зависимых источников с дополнительной информацией на декодере. Далее как в тексте диссертационной работы, так и в тексте автореферата, где это не будет вызывать неоднозначности, будем называть подобные системы системами распределенного кодирования видеоданных, а кодеки – распределенными видеокодеками.

В схемах сжатия, основанных на распределенном кодировании, устранение временной избыточности осуществляется с помощью межкадрового предсказания, выполняемого на стороне декодера, что позволяет существенно снизить сложность обработки на передатчике. С точки зрения распределенного кодирования результат предсказания является дополнительной информацией, а саму процедуру межкадрового предсказания называют генерацией дополнительной информации. Кодеру в подобных системах необходимо сформировать уточняющую информацию, которая позволит декодеру исправить ошибки межкадрового предсказания и тем самым восстановить кадр. Существующие реализации распределенных видеокодеков используют для исправления ошибок межкадрового предсказания методы помехоустойчивого кодирования. Описанную процедуру обработки видеоданных принято интерпретировать с использованием понятия *виртуального канала*, по которому «передаются» исходные данные, а на выходе наблюдается результат межкадрового предсказания. Чтобы исправить возникшие «ошибки», на передатчике располагается помехоустойчивый кодер, который интерпретирует данные дополнительной информации как информационную последовательность и формирует по ней проверочные символы, используя заранее определенный помехоустойчивый код. На приемнике помехоустойчивый декодер объединяет полученные от кодера проверочные символы с результатом предсказания, после чего применяет процедуру декодирования, исправляя таким образом ошибки предсказания. Сжатие достигается в том случае, если объем переданных помехоустойчивому декодеру проверочных данных меньше, чем объем исходных данных.

После описания ключевых идей и понятий распределенного кодирования в первом разделе для нескольких простых источников приведен ряд воз-

можных схем сжатия, для которых показано, что распределенное кодирование позволяет осуществлять сжатие со скоростью, равной энтропии этих источников.

Далее в разделе рассматриваются практические подходы к распределенному кодированию видеоданных. Приведена классификация моделей распределенных видеокодеков и обоснован выбор для дальнейших исследований модели, основанной на обработке целых кадров (концепция Стэнфорд) в спектральной области. Для выбранной модели приведено описание эталонной модели распределенного кодирования DISCOVER, в рамках которой выделены ключевые операции, влияющие на степень сжатия: генерация дополнительной информации и оценка параметров ошибок межкадрового предсказания. Исследованию этих операций посвящены следующие разделы диссертационной работы.

Во *втором разделе* исследованы способы генерации дополнительной информации для систем распределенного кодирования видеоданных. Наиболее распространенным подходом является интерполяция промежуточного кадра на стороне декодера с использованием уже восстановленных кадров. Качество интерполяции во многом определяется точностью процедуры оценки движения, в результате которой необходимо получить знание о действительном смещении объектов между соседними кадрами, или *истинном движении*. В связи с этим в данном разделе основное внимание уделено анализу моделей истинного движения, а также разработке алгоритмов интерполяции, основанных на оценке истинного движения.

Здесь и далее будем рассматривать процесс сжатия видеопоследовательности на примере двух смежных опорных кадров F_p и F_f , а также одного промежуточного кадра F_i . Одним из ключевых допущений при формализации понятия истинного движения является допущение о гладкости, гласящее, что векторное поле, соответствующее истинному движению, является кусочно-гладким. Под гладкостью понимается коррелированность векторов движения для пикселей, принадлежащих одному движущемуся объекту. При этом на границах объектов в векторном поле наблюдаются резкие перепады. Будем считать, что векторное поле, соответствующее истинному движению, может быть получено

в результате решения следующей оптимизационной задачи:

$$\mathbf{V}^* = \arg \min_{\mathbf{V} \in \mathcal{V}} E(\mathbf{F}_p, \mathbf{F}_f, \mathbf{V}, \alpha), \quad (1)$$

где \mathcal{V} - это множество всех возможных векторных полей \mathbf{V} между кадрами \mathbf{F}_p и \mathbf{F}_f и

$$E(\mathbf{F}_p, \mathbf{F}_f, \mathbf{V}, \alpha) = E_1(\mathbf{F}_p, \mathbf{F}_f, \mathbf{V}) + \alpha E_2(\mathbf{V}).$$

Слагаемое $E_1(\mathbf{F}_p, \mathbf{F}_f, \mathbf{V})$ соответствует энергии разностного кадра, полученного с использованием векторного поля \mathbf{V} ; слагаемое $E_2(\mathbf{V})$ отражает гладкость поля \mathbf{V} ; $\alpha \geq 0$ – коэффициент регуляризации между энергией разностного кадра и гладкостью векторного поля.

Следует отметить, что с помощью формулы (1) описывается только модель движения, и векторные поля, соответствующие локальным и глобальным минимумам этой модели, не обязательно точно отражают истинное движение объектов на кадрах. В частности, модель не учитывает неравномерность движения на границах объектов, где векторное поле, как правило, не является гладким. Поэтому найденное для данной модели решение не гарантирует оптимального визуального качества интерполяции. В связи с этим более хорошие результаты с точки зрения визуального качества достигаются, как правило, с использованием подоптимальных алгоритмов, рассматривающих оптимизацию (1) неявно. Такие алгоритмы, с одной стороны, основаны на эвристических подходах, нацеленных на повышение визуального качества интерполированных кадров, а с другой стороны, неявно косвенно минимизируют (1), т.е. выдают результат, согласованный с моделью истинного движения. В данном разделе приводится описание нового алгоритма оценки движения, принадлежащего именно к такому классу алгоритмов.

В основе разработанного эвристического алгоритма лежит многоуровневая иерархическая билатеральная процедура блоковой оценки движения с дополнительным итеративным поиском. Билатеральный подход в отличие от однонаправленного поиска, позволяет найти для позиции каждого блока на промежуточном кадре вектор движения. На каждом уровне иерархии выполняется оценка движения для блоков фиксированного размера. Здесь и далее будем называть процесс обработки одного уровня иерархии стадией. Выходом стадии является векторное поле \mathbf{V}^* , доставляющее локальный минимум (1). Получен-

ные векторы движения используются в качестве начального смещения на следующей стадии. Иерархическая оценка реализована с использованием много-сеточного подхода, в котором размер блока уменьшается с номером стадии. На каждой стадии для фиксированного размера блока выполняются две основные операции (рис. 1):

1. начальная билатеральная оценка движения, используемая для поиска предварительного векторного поля;
2. итеративный дополнительный поиск с вычислением надежностей векторов, повышающий гладкость векторного поля.

Начальная билатеральная оценка движения является первой процедурой поиска векторов на стадии, при этом все блоки обрабатываются независимо друг от друга. Для каждого блока выполняется процедура локально-оптимального поиска вектора движения с использованием подхода, основанного на градиентном спуске. После завершения начальной оценки каждому блоку на промежуточном кадре поставлен в соответствие вектор движения, минимизирующий ошибку сопоставления блоков. Затем применяется дополнительный поиск для повышения гладкости полученного векторного поля. Процедура дополнительного поиска реализована как итеративный поиск, основанный на предсказании с учетом надежности векторов. Вектор считается надежным, если соответствующая ему билатеральная ошибка сопоставления мала и его значение коррелировано с соседями. Процесс предсказания оперирует с надежными векторами из множества соседей текущего блока на данном уровне иерархии и с надежными векторами на предыдущих уровнях иерархии. Эти векторы формируют множество кандидатов для интерполяции текущего блока. Кандидаты используются для инициализации градиентного спуска при поиске оптимального вектора для текущего блока.

Далее в данном разделе приводится описание метода аппроксимации промежуточного с использованием процедуры временной интерполяции, в основе которой лежит разработанный алгоритм оценки движения. Для повышения визуального качества аппроксимирующих кадров в процедуре интерполяции используется ряд эвристических подходов:

- определение статичных регионов, т.е. таких регионов, интенсивность пикселей в которых не меняется между кадрами;

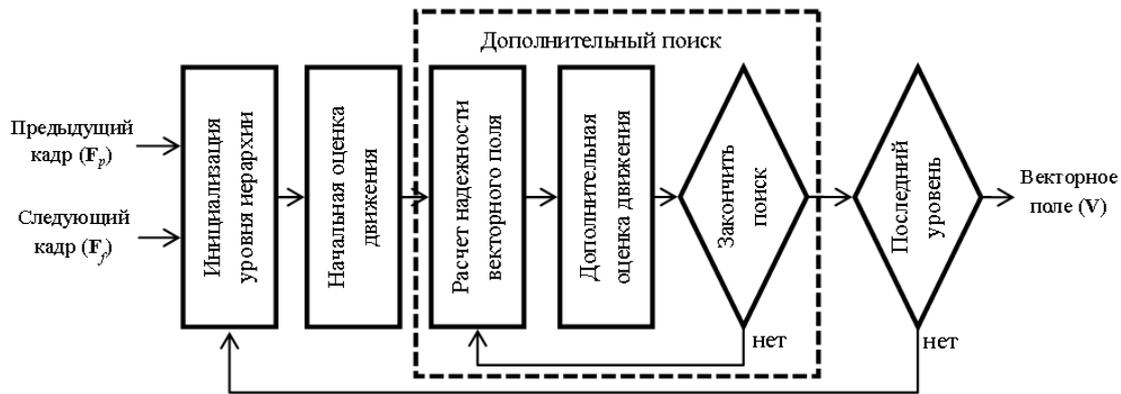


Рис. 1 — Последовательность действий на одной стадии разработанного алгоритма оценки движения.

- интерполяция кадра с использованием компенсации движения с перекрытиями для уменьшения блочности на аппроксимирующем кадре.

Алгоритм 1. Алгоритм межкадрового предсказания.

Входные данные: пара смежных восстановленных опорных кадров F_p и F_f .

Выходные данные: промежуточный кадр F_i .

1. Определение статичных регионов на кадрах F_p и F_f .
2. Выполнение оценки движения между кадрами F_p и F_f с помощью алгоритма оценки истинного движения (рис. 1).
3. Расчет промежуточного кадра F_i с использованием компенсации движения с перекрытиями.
4. Наложение статичных регионов на кадр F_i .

Данный алгоритм предлагается использовать как составную часть процедуры аппроксимации кадра на стороне декодера, что позволит уменьшить количество ошибок при генерации дополнительной информации.

В *третьем разделе* рассматривается задача оценки параметров ошибок межкадрового предсказания, на основании которых в дальнейшем формируется мягкий вход помехоустойчивого декодера, выполняющего исправление искажений в кадре дополнительной информации.

Декодер выполняет блоковое спектральное преобразование и квантование для пикселей кадра F_a . Для каждой группы спектральных коэффициентов формируется набор индексов квантов $\hat{\mathbf{q}}^{(b)} = (\hat{q}_1^{(b)}, \hat{q}_2^{(b)}, \dots, \hat{q}_m^{(b)})$, где b – индекс

группы спектральных коэффициентов, m – число спектральных коэффициентов в группе. Отметим, что кодер, выполнив аналогичные операции, получает набор индексов квантов $\mathbf{q}^{(b)} = (q_1^{(b)}, q_2^{(b)}, \dots, q_m^{(b)})$ для оригинального промежуточного кадра \mathbf{F}_i . Ошибка межкадрового предсказания в спектральной области определяется как

$$\mathbf{n}^{(b)} = \mathbf{q}^{(b)} - \hat{\mathbf{q}}^{(b)}.$$

Далее в разделе рассматриваются общепринятые допущения, используемые при разработке алгоритмов оценки параметров ошибок межкадрового предсказания.

Допущение 1. Для каждой группы спектральных коэффициентов декодер может выполнить аппроксимацию ошибок $\hat{\mathbf{n}}^{(b)} = (\hat{n}_1^{(b)}, \hat{n}_2^{(b)}, \dots, \hat{n}_m^{(b)})$, причем $\hat{n}_i^{(b)} = n_i^{(b)} + e_i^{(b)}$, где $e_i^{(b)}$ – случайное искажение, имеющее некоторый закон распределения с нулевым математическим ожиданием и конечной дисперсией.

Допущение 2. Виртуальный канал является каналом с аддитивным нестационарным Лапласовым шумом, т. е. каждое значение $n_i^{(b)}$ является реализацией случайной величины $N_i^{(b)}$, имеющей закон распределения Лапласа с нулевым математическим ожиданием:

$$N_i^{(b)} \sim \frac{\alpha_i^{(b)}}{2} e^{-\alpha_i^{(b)} |x|}.$$

Задача оценки параметров ошибок межкадрового предсказания заключается в поиске для каждого квантованного спектрального коэффициента параметра масштаба распределения Лапласа. Для решения этой задачи в DISCOVER применяется эвристическая процедура, которую здесь и далее будем называть базовым алгоритмом оценки параметров ошибок межкадрового предсказания. В базовом алгоритме для каждой группы спектральных коэффициентов независимо выполняются три основных действия.

1. Оценка параметра масштаба с использованием всех коэффициентов в группе спектральных коэффициентов (глобальный параметр).
2. Оценка параметра масштаба для каждого коэффициента независимо (локальный параметр).

3. Выбор для каждого коэффициента либо локального, либо глобального параметра.

Данный алгоритм не учитывает свойство пространственной группировки ошибок интерполяции с большой амплитудой, особенно характерное для последовательностей с быстрым движением. Чтобы обойти данный недостаток, в работе предлагается модификация базового алгоритма, заключающаяся в предварительном разбиении аппроксимации ошибок на подмножества с однородными характеристиками. Для учета возможных ошибок аппроксимации предлагается воспользоваться пространственной зависимостью между ошибками и отфильтровать индексы подмножеств с помощью двумерного медианного фильтра. Затем в каждом подмножестве независимо применяется базовый алгоритм.

Алгоритм 2. Алгоритм оценки параметров ошибок в виртуальном канале.

Входные данные: аппроксимация ошибки межкадрового предсказания в спектральной области $\hat{n}^{(b)}$, количество сегментов $s^{(b)}$.

Выходные данные: параметр масштаба распределения Лапласа, рассчитанный для каждого спектрального коэффициента.

1. Выполнение разбиения $\tilde{n}^{(b)}$ на $s^{(b)}$ подмножеств.
2. Двумерная медианная фильтрация карты подмножеств.
3. Оценка параметров с использованием базового алгоритма в каждом подмножестве независимо.

В завершении третьего раздела рассматриваются вопросы, связанные с формализацией задачи оценки параметров ошибок межкадрового предсказания. Отмечается, что в настоящее время не существует строгого математического описания модели ошибок в виртуальном канале, что усложняет анализ существующих алгоритмов оценки параметров ошибок межкадрового предсказания и ухудшает понимание специфики и особенностей данной задачи в целом. Предлагается следующий расширенный набор допущений. Во введенном ранее допущении 2 уберем ограничение, связанное с тем, что рассматриваются только распределения Лапласа.

Допущение 2 (). Виртуальный канал является каналом с аддитивным нестационарным шумом.*

Для определенности будем полагать, что ошибки в виртуальном канале являются случайными величинами, закон распределения каждой из которых принадлежит множеству $\mathcal{D}^{(b)} = \{D_i^{(b)}(x|\boldsymbol{\theta}_i^{(b)})\}$, $i = \overline{1, k^{(b)}}$, где $D_i^{(b)}(x|\boldsymbol{\theta}_i^{(b)})$ задает вероятностный закон и его параметры для распределения с индексом i .

Будем называть множество случайных величин $\mathcal{N}^{(b)} = \{N_1^{(b)}, N_2^{(b)}, \dots, N_m^{(b)}\}$ *полем ошибок*, а множество случайных величин $\mathcal{A}^{(b)} = \{A_1^{(b)}, A_2^{(b)}, \dots, A_m^{(b)}\}$, где $A_j^{(b)} \in \{1, 2, \dots, k^{(b)}\}$ – *полем распределений*.

В процессе моделирования виртуального канала каждой случайной величине из $\mathcal{N}^{(b)}$ должно быть сопоставлено распределение из $\mathcal{D}^{(b)}$. Будем обозначать событие, что $N_i^{(b)}$ имеет закон распределения $D_j^{(b)}(x|\boldsymbol{\theta}_j^{(b)})$, как $N_i^{(b)} \sim D_j^{(b)}(x|\boldsymbol{\theta}_j^{(b)})$. Событие, заключающееся в сопоставлении законов распределений из $\mathcal{D}^{(b)}$ со случайными величинами из $\mathcal{N}^{(b)}$, можно рассматривать как реализацию поля распределений $\mathcal{A}^{(b)}$.

Допущение 3. *Поле распределений является Марковским случайным полем.*

Использование данных допущений позволяет сформулировать следующую целевую функцию для задачи оценки параметров ошибок межкадрового предсказания:

$$\mathbf{a}^{(b)*} = \arg \max_{\mathbf{a} \in \{1, 2, \dots, k^{(b)}\}^m} \prod_{i=1}^m p(\hat{n}_i^{(b)} | a_i) \exp \left(- \sum_{c \in \mathcal{C}} \psi_c(\mathbf{a}) \right) \quad (2)$$

где $\mathbf{a}^{(b)*}$ – реализация поля распределений, соответствующая сопоставлению законов $\mathcal{D}^{(b)}$ со случайными величинами из $\mathcal{N}^{(b)}$, максимизирующему правдоподобие аппроксимации шума; \mathcal{C} – множество клик графа, описывающего Марковское свойство поля распределений; $\psi_c(\mathbf{a})$ – потенциалы, задающие совместные функции распределений на зависимых случайных величинах в поле распределений.

В работе показано, что в результате применения модифицированного алгоритма получается более высокое значение целевой функции по сравнению с базовым алгоритмом (2).

Четвертый раздел посвящен вопросам, связанным с проведением сравнительного анализа алгоритмов, разработанных в диссертационной работе. В начале раздела приведено детальное описание и результаты верификации про-

граммной модели распределенного видеокодека, на базе которой далее проводились все эксперименты. В основе программной модели лежат алгоритмы сжатия, используемых в DISCOVER. Также в начале раздела дано описание сформированного тестового множества видеопоследовательностей, указаны критерии сравнения.

Далее в разделе перечислены основные вопросы, которые должны быть рассмотрены при проведении сравнительного анализа алгоритмов обработки информации в рамках распределенного кодирования. В частности отмечено, что при анализе алгоритмов межкадрового предсказания необходимо учесть тот факт, что модуль генерации дополнительной информации находится в тесной взаимосвязи с прочими модулями кодека, что делает не совсем корректным использование полной схемы кодирования для сравнения различных алгоритмов генерации дополнительной информации. В диссертационной работе для решения данной проблемы предлагается использовать следующий подход. Реализуется упрощенная схема распределенного видеокодека без обратной связи, т. е. передача проверочной информации декодеру от кодера не осуществляется. Отсутствие обратной связи позволяет осуществлять сравнение алгоритмов генерации дополнительной информации в одинаковых условиях, т. е. без влияния прочих операций, выполняемых декодером. В подобной схеме битовые затраты на передачу промежуточных кадров составляют 0 бит, а в качестве восстановленных кадров выдаются аппроксимирующие кадры, полученные в процессе генерации дополнительной информации. Таким образом суммарные битовые затраты на сжатие всей последовательности определяются только затратами на хранение информации об опорных кадрах. Изменяя степень сжатия опорных кадров строятся кривые скорость-искажение, по которым осуществляется сравнительная оценка различных методов генерации дополнительной информации.

Данная схема была реализована в рамках разработанного программного комплекса и с её использованием было показано, что предложенный алгоритм межкадрового предсказания существенно превосходит базовый алгоритм из кодека DISCOVER. В частности, битовые затраты на последовательностях с быстрым движением снижаются на 10 – 30%, на последовательностях с медленным движением – на 3 – 10%. Далее было продемонстрировано, что применение модифицированного алгоритма моделирования виртуального канала позволя-

ет снизить битовые затраты на последовательностях с быстрым движением на 0.1 – 1%.

В конце раздела приведены результаты сравнения реализованной программной модели распределенного видеокodeка, включающей все предложенные алгоритмы, с существующими алгоритмами сжатия. Показано, что на всех последовательностях из тестового множества разработанная программная модель выигрывает по степени сжатия у codeка DISCOVER 4 – 16%. При этом оба codeка проигрывают на всех последовательностях codeку H.264 Inter. На последовательностях с высокой сложностью движения (Football и Soccer) и DISCOVER, и разработанный алгоритм проигрывают как H.264 Inter, так и H.264 Intra. Однако следует отметить, что на высоком качестве для последовательности Soccer разработанный алгоритм, в отличие от DISCOVER, показывает качество, близкое к H.264 Intra.

В *заключении* приведены основные результаты, полученные в диссертационной работе.

Основные результаты работы

1. Разработан алгоритм межкадрового предсказания для codeков видеоинформации, основанных на принципах кодирования зависимых источников с дополнительной информацией на декодере, использующий метод временной интерполяции, позволяющий за счет использования процедуры оценки истинного движения объектов уменьшить по сравнению с существующими алгоритмами число ошибок интерполяции и тем самым повысить степень сжатия на 10 – 30% для последовательностей с быстрым движением и на 3 – 10% на последовательностях с медленным движением.
2. Предложен метод выполнения сравнительной оценки алгоритмов межкадрового предсказания в системах кодирования зависимых источников видеоинформации с дополнительной информацией на декодере, учитывающий тот факт, что модули системы распределенного кодирования оказывают сильное влияние друг на друга. В рамках предложенного метода разработана упрощенная модель видеокodeка без обратной связи, позволяющая выполнять сравнительный анализ

алгоритмов межкадрового предсказания, устраняя влияние остальных модулей системы.

3. Предложена модификация базового алгоритма оценки параметров ошибок межкадрового предсказания из кодека DISCOVER, основанная на процедуре разбиения неоднородного множества ошибок на подмножества с однородными характеристиками с последующим применением базового алгоритма в каждом подмножестве независимо. Использование модифицированного алгоритма позволяет повысить на последовательностях с быстрым движением степень сжатия на 0.6 – 1% по сравнению с базовым алгоритмом.
4. Проведено исследование существующих моделей виртуального канала в системах распределенного кодирования видеоданных. Показано, что общепринятый набор допущений не в полной мере описывает особенности ошибок межкадрового предсказания. Предложен новый расширенный набор допущений и на его основе предложена новая модель виртуального канала, основанная на модели ошибок, представляющей процесс возникновения искажений с помощью скрытого Марковского случайного поля (СМСП).

Основное содержание работы изложено в следующих публикациях (статьи 1-2 опубликованы в изданиях, включенных в перечень ВАК, статьи 3-7 индексируются в Scopus):

Публикации автора по теме диссертации

1. *Веселов, А.И.* Метод генерации сторонней информации для систем распределенного кодирования видеоисточников / А.И. Веселов, М.Р. Гильмутдинов, Б.С. Филиппов // *Известия вузов. Приборостроение*. — 2013. — Т. 56, № 8. — С. 62–68.
2. *Веселов, А.И.* Алгоритм временной интерполяции кадров, основанный на процедуре итеративной оценки движения / А.И. Веселов, М.Р. Гильмутдинов // *Информационно-управляющие системы*. — 2014. — Т. 71, № 4. — С. 25–33.

3. Scalable-to-lossless transform domain distributed video coding / A. Veselov, X. Huang, A. Ukhanova et al. // IEEE International Workshop on Multimedia Signal Processing. — 2010. — Oct. — Pp. 327–332.
4. *Veselov, A.* Analysis of Side Information Generation Impact on Distributed Video Coding Performance / A. Veselov, B. Filippov, M. Gilmudinov // XIII International Symposium on Problems of Redundancy in Information and Control Systems. — 2012. — Sept. — Pp. 26–30.
5. *Veselov, A.* Iterative hierarchical true motion estimation for temporal frame interpolation / A. Veselov, M. Gilmudinov // IEEE 16th International Workshop on Multimedia Signal Processing. — 2014. — Sept. — Pp. 1–6.
6. *Veselov, A.* A Generalized Correlation Noise Model for Pixel Domain Wyner-Ziv Video Coding / A. Veselov, M. Gilmudinov // 6th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops. — 2014. — Oct.
7. Non-stationary Correlation Noise Modeling for Transform Domain Wyner-Ziv Video Coding / A. Veselov, B. Filippov, V. Yastrebov, M. Gilmudinov // Smart Innovation, Systems and Technologies. — Springer, 2015. — Vol. 40. — Pp. 179–189.
8. *Веселов, А. И.* Количественное оценивание визуальных искажений возникающих при кодировании с использованием дискретного косинусного преобразования / А. И. Веселов, М. Р. Гильмутдинов // *Вопросы передачи и защиты информации.* — 2011. — С. 32–50.
9. *Веселов, А. И.* Оценка энтропии одного частного случая скрытого марковского процесса / А. И. Веселов // Научная сессия ГУАП. — 2010.
10. *Веселов, А.И.* Анализ моделей шума в виртуальном канале в системах распределенного кодирования видеоданных / А.И. Веселов // Научная сессия ГУАП. — 2011. — С. 48–51.
11. Complexity Analysis of Adaptive Binary Arithmetic Coding Software Implementations / E. Belyaev, A. Veselov, A. Turlikov, L. Kai // Lecture Notes in Computer Science. — Springer, 2011. — Vol. 6869. — Pp. 598–607.

12. *Веселов, А.И.* Обработка видеoinформации в системах сжатия, основанных на принципах кодирования зависимых источников / А.И. Веселов, М.Р. Гильмутдинов. — ГУАП, 2014. — С. 71.
13. Пат. **WO2013095180 A1**, Complexity scalable frame rate up-conversion / Веселов А.И., Гильмутдинов М.Р.; приоритетная заявка РСТ/RU2011/001020; заявл. 22.12.2011; опубл. 27.06.2013 (США).
14. Пат. **WO2013100791 A1**, Method of and apparatus for scalable frame rate up-conversion / Веселов А.И., Гильмутдинов М.Р.; приоритетная заявка РСТ/RU2011/001059; заявл. 30.12.2011; опубл. 04.07.2013 (США).
15. Пат. **US20140002732 A1**, Method and system for temporal frame interpolation with static regions excluding / Веселов А.И., Гильмутдинов М.Р.; приоритетная заявка US 13/539,035; заявл. 29.06.2012; опубл. 02.01.2014 (США).