

На правах рукописи

Беляев Евгений Александрович

**УПРАВЛЕНИЕ ПАРАМЕТРАМИ АЛГОРИТМА СЖАТИЯ
ВИДЕОИНФОРМАЦИИ ПРИ ПЕРЕДАЧЕ ДАННЫХ
В СИСТЕМАХ МОБИЛЬНОЙ СВЯЗИ**

Специальность: 05.13.01

«Системный анализ, управление и обработка информации
(в технике и технологиях)»

Автореферат диссертации на соискание ученой степени
кандидата технических наук

Санкт-Петербург
2008

Работа выполнена на кафедре безопасности информационных систем Государственного образовательного учреждения высшего профессионального образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения» (СПГУАП).

Научный руководитель:

кандидат технических наук, доцент Тюрликов Андрей Михайлович

Официальные оппоненты:

доктор технических наук, профессор Красильников Николай Николаевич

кандидат технических наук, доцент Украинский Олег Владимирович

Ведущая организация:

ФГУП «Ленинградский отраслевой научно-исследовательский институт связи»

Защита состоится «___» _____ 2009 г. в ____ час. на заседании диссертационного совета Д212.233.02 при Государственном образовательном учреждении высшего профессионального образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения» по адресу: 190000, г. Санкт-Петербург, ул. Большая Морская, д.67.

С диссертацией можно ознакомиться в библиотеке СПГУАП.

Автореферат разослан «___» _____ 2008 г.

Ученый секретарь диссертационного совета

Л.А. Осипов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. В связи с бурным ростом инфокоммуникационных технологий все большую роль в общем информационном потоке играют мультимедийные данные. При этом существенную долю передаваемых по проводным и беспроводным каналам данных составляют видеоданные. Так как в качестве приемников и передатчиков видеoinформации все чаще используются мобильные устройства, актуальна задача разработки алгоритмов сжатия и передачи видеoinформации с учетом специфики этих устройств.

В области разработки алгоритмов сжатия и передачи видеоданных в системах мобильной связи можно выделить два основных направления исследований: передача по низкоскоростным каналам связи на большие расстояния (например, в системах цифрового телевизионного вещания DVB-H, технологии передачи данных EGPRS для мобильных сетей GSM, технологии WiMAX и др.) и передача видеоданных высокого разрешения с малой задержкой по высокоскоростным каналам связи на небольшие расстояния (например, в системах, основанных на технологии WiFi, UWB и др.). Для первого направления характерна задача разработки алгоритмов, направленных на дальнейшее увеличение степени сжатия передаваемых видеоданных. Для второго направления характерна задача разработки алгоритмов, учитывающих ограничения на вычислительную сложность и объем памяти мобильных устройств.

Основной целью работы является исследование и разработка методов управления параметрами алгоритма сжатия видеoinформации, направленных на повышение эффективности передачи видеоданных в системах мобильной связи.

Для достижения поставленной цели были исследованы следующие задачи:

1. Задача повышения эффективности адаптивного арифметического кодирования при сжатии видеoinформации.
2. Задача устранения межкадровой (временной) избыточности видеoinформации для случая передачи по каналам связи с низкой пропускной способностью.
3. Задача управления скоростью кодирования видеoinформации при передаче данных в системах мобильной связи.
4. Задача разработки алгоритмов с учетом особенностей реализации системы сжатия и передачи видеoinформации, в частности, с учетом ограничения на объем памяти кодирующего и декодирующего устройств.

Объектом исследования является система сжатия и передачи видеoinформации.

Предметом исследования являются методы управления параметрами алгоритма сжатия видеoinформации.

Методологическую и теоретическую основу исследования составили научные труды отечественных и зарубежных ученых как в области теории информации (В.Д. Колесник, Б.Д. Кудряшов, Р.Е. Кричевский, Б.Я. Рябко, В.К. Трофимов, Ю.М. Штарьков, Р. Галлагер, Т. Лейтон, Г. Лэнгдон, Р. Ривест и др.), так и в области обработки и сжатия видеoinформации (Н.Н. Красильников, И. Добеши, А. Катсаггелос, Д. Марпе, Г. Шустер и др.).

Методы исследования. Для решения поставленных задач использованы методы цифровой обработки сигналов, теории информации и теории оптимизации.

Для получения практических результатов были использованы открытые для общего использования кодеки, поддерживающие стандарты сжатия видеoinформации MPEG-2, H.264/AVC и JPEG2000. Реализация предложенных алгоритмов была осуществлена на языке программирования Си в среде Microsoft Visual Studio 2005.

Научная новизна диссертационной работы заключается в следующем.

1. Предложен алгоритм «виртуального скользящего окна», который может быть использован для повышения эффективности адаптивного арифметического кодирования при сжатии видеoinформации.
2. Предложена модификация алгоритма «иерархической оценки движения», которая позволяет повысить эффективность кодера видеoinформации при передаче данных по каналам связи с низкой пропускной способностью.
3. Впервые сформулирована задача управления скоростью кодирования для группы видеисточников по минимаксному критерию искажения. Предложено решение задачи на основе модифицированного алгоритма «иерархической оценки движения».
4. Сформулирована задача управления скоростью кодирования по минимаксному критерию искажения при передаче видеoinформации по каналу связи при ограничении на объем памяти кодирующего и декодирующего устройств. Предложены алгоритмы управления, позволяющие повысить эффективность кодирования для видеопоследовательностей, содержащих фрагменты компьютерной графики.

Практическая значимость исследования. Результаты диссертационной работы могут быть использованы при проектировании систем сжатия и передачи видеoinформации, а также для разработки стандартов сжатия видеoinформации.

Алгоритм «виртуального скользящего окна», модифицированный алгоритм «иерархической оценки движения» и алгоритм управления скоростью кодирования для группы видеисточников могут быть использованы для повышения эффективности передачи видеоданных по каналам связи с низкой пропускной способностью. Алгоритм управления скоростью кодирования видеoinформации при ограничении на объем памяти кодирующего и декодирующего устройств может быть применен для для повышения эффективности передачи видеоданных на небольшие расстояния по высокоскоростным каналам связи с малой задержкой передачи данных.

Апробация результатов исследования. Основные результаты диссертационной работы докладывались на семинарах кафедры информационных систем и кафедры безопасности информационных систем СПГУАП, а также на следующих научно-технических конференциях: 8я, 9я и 10я научная сессия СПГУАП, «2006 IEEE 10th International Symposium on Consumer Electronics», «XI International Symposium on Problems of Redundancy in Information and Control Systems», «15-я Международная научно-техническая конференция «Проблемы передачи и обработки информации в сетях и системах

телекоммуникаций», «The 15-th International Conference on Communications», «The 11th International Symposium on Wireless Personal Multimedia Communications».

Внедрение результатов исследования. Результаты диссертационной работы используются в учебном процессе кафедры безопасности информационных систем СПГУАП. Разработанный алгоритм управления скоростью кодирования видеоинформации с учетом ограничения на объем памяти кодирующего и декодирующего устройств был использован в рамках проекта «Беспроводной экран» в ЗАО «Интел А/О».

Публикации по теме диссертационной работы. Основные результаты работы изложены в 12-ти печатных работах, из которых 3 опубликованы в научных журналах, входящих перечень ВАК.

Основные положения, выносимые на защиту:

1. Алгоритм «виртуального скользящего окна» для повышения эффективности адаптивного арифметического кодирования при сжатии видеоинформации.
2. Модифицированный алгоритм «иерархической оценки движения» для устранения межкадровой (временной) избыточности видеоинформации.
3. Метод управления скоростью кодирования видеоинформации при передаче данных в системах мобильной связи.

Структура и объем работы. Диссертационная работа состоит из введения, четырех разделов, заключения, списка литературы и четырех приложений. Работа содержит 172 страницы машинописного текста, включая 55 рисунков. Список литературы содержит 82 источника.

СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обоснована актуальность темы диссертации, сформулированы цели диссертационной работы, основные задачи, основные положения, выносимые на защиту, а также кратко изложено содержание работы по разделам.

В **первом** разделе приводится структурная схема кодера видеоинформации.

Основными характеристиками кодера видеоинформации являются *скорость кодирования* (количество бит, формируемых кодером в единицу времени) и *уровень искажения* восстановленной видеопоследовательности относительно исходной. В качестве мер искажения в диссертационной работе используются среднеквадратическая ошибка и пиковое отношение сигнал/шум (Peak signal-to-noise ratio, PSNR).

Кодер видеоинформации выполняет следующие основные операции:

1. *Устранение статистической избыточности источника видеоинформации.*
2. *Квантование.*
3. *Статистическое сжатие без потерь.*
4. *Управление скоростью кодирования источника видеоинформации в соответствии с заданным критерием искажения.*

Избыточность источников видеoinформации подразделяют на пространственную и временную. Пространственная избыточность обусловлена статистической зависимостью значений яркости между соседними отсчетами внутри видеокадра. Для устранения этого вида избыточности наиболее часто используется дискретное косинусное преобразование и дискретное вейвлетное преобразование. Временная избыточность связана со статистической зависимостью значений яркости между соответствующими отсчетами соседних видеокадров. Эту зависимость обычно устраняют, используя блоковую компенсацию движения. В этом случае видеокадр разбивается на прямоугольные блоки. Затем выполняется процедура *оценки движения*, при которой в текущем кадре для каждого блока осуществляется поиск в базовом кадре наиболее «похожего» блока, смещенного на вектор, называемый *вектором движения*.

Для регулирования скорости кодирования видеoinформации применяется *квантование*, в результате которого вносится искажение.

Для *статистического сжатия без потерь* векторов движения и данных, полученных после квантования, используются известные в теории информации подходы. Наиболее часто применяются побуквенные коды переменной длины (например, код Хаффмана) или адаптивное арифметическое кодирование.

Параметры алгоритма сжатия влияют на соотношение скорости кодирования и уровня искажения. Например, при увеличении шага квантования как правило повышается уровень искажения и уменьшается скорость кодирования. При передаче видеoinформации задается некоторый набор ограничений, с учетом которого необходимо найти «оптимальное» по заданному критерию искажения управление параметрами кодера видеoinформации. Алгоритмы, которые решают такую задачу, называются *алгоритмами управления скоростью кодирования видеoinформации*. В качестве ограничений могут выступать пропускная способность канала связи, задержка передачи данных, а также особенности реализации системы сжатия и передачи, например, ограничение на вычислительную сложность или объем памяти кодирующего и декодирующего устройств.

В первом разделе выделены особенности передачи данных от мобильных устройств, описаны известные алгоритмы сжатия данных без потерь и с потерями, а также алгоритмы устранения пространственной и временной избыточности видеoinформации.

Второй раздел посвящен задаче повышения эффективности передачи видеоданных по каналам связи с низкой пропускной способностью при помощи усовершенствования адаптивного арифметического кодирования.

Для статистического сжатия без потерь в современных стандартах сжатия видеoinформации (например, H.264/AVC и JPEG2000) используется контекстный адаптивный двоичный арифметический кодер. Перед арифметическим кодированием недвоичные данные сначала преобразуются в двоичную последовательность, которая при помощи контекстного моделирования разбивается на множество двоичных подпоследовательностей. Затем оценивается вероятность появления единицы в двоичной подпоследовательности, соответствующей каждому контексту.

Двоичную подпоследовательность можно разбить на интервалы так, что статистические характеристики данных внутри каждого интервала постоянны, но могут существенно отличаться для разных интервалов. Достигаемая при

арифметическом кодировании степень сжатия таких данных во многом зависит как от точности оцениваемой внутри каждого интервала статистики, так и от скорости адаптации к изменяющейся между интервалами статистике. Поэтому алгоритмы оценки вероятности должны сочетать в себе как точность оценки, так и скорость адаптации.

Описанные выше свойства могут быть достигнуты при использовании конструкции «скользящего окна», в котором вероятность появления очередного символа x_{t+1} источника определяется в результате анализа содержимого окна, то есть последовательности символов $x_{t-W+1}, x_{t-W+2}, \dots, x_t$, где $W \geq 1$ - длина окна. После кодирования очередного символа содержимое окна сдвигается на одну позицию, новый символ x_{t+1} заносится в освободившуюся ячейку, а последний символ x_{t-W+1} удаляется. Оценка вероятности появления единицы на выходе двоичного источника при использовании «скользящего окна» определяется при помощи оценки Кричевского-Трофимова: $\hat{p}_{t+1} = \frac{n_t^1 + 1/2}{W+1}$, где n_t^1 - число единиц в окне. Основным недостатком данного подхода является необходимость хранения в памяти кодера и декодера W последних символов для каждого контекста.

Для устранения данного недостатка может быть использована аппроксимация «скользящего окна», при которой в памяти кодера и декодера хранится только число символов в окне, которое вычисляется по некоторому правилу. Одно из таких правил, названное «мнимым скользящим окном», предложено Б.Я. Рябко. Согласно этому правилу, после кодирования символа x_t из окна удаляется не последний, а случайный символ y_t^e и число единиц в окне $n_{t+1}^1 = n_t^1 - y_t^e + x_t$, где $y_t^e \in \{0, 1\}$ - случайная величина, которая принимает значение, равное единице с вероятностью $\frac{n_t^1}{W}$ и нулю с вероятностью $1 - \frac{n_t^1}{W}$. Аналогичным образом генерируется случайная величина y_t^d после декодирования символа x_t . Причем для однозначного декодирования значения генерируемых случайных величин должны совпадать $y_t^e = y_t^d$, то есть на стороне кодера и декодера должна использоваться специальным образом сформированная псевдослучайная последовательность, что усложняет реализацию алгоритма.

В диссертационной работе показано, как можно избежать генерирования случайной величины. Для этого случайная величина y_t заменяется ее математическим ожиданием. При этом правило пересчета числа единиц в окне изменяется следующим образом: $n_{t+1}^1 = (1 - \frac{1}{W})n_t^1 + x_t$.

При таком подходе возникают операции с вещественными числами. Умножив обе части полученного равенства на множитель c можно перейти к целочисленной реализации правила пересчета, которое в диссертационной работе названо «виртуальным скользящим окном». Согласно данному правилу число единиц s_{t+1} в окне из cW ячеек после кодирования очередного символа x_t пересчитывается по следующему правилу:

$$s_{t+1} = \begin{cases} s_t + \left\lfloor \frac{cW - s_t + \frac{W}{2}}{W} \right\rfloor, & \text{при } x_t = 1 \\ s_t - \left\lfloor \frac{s_t + \frac{W}{2}}{W} \right\rfloor, & \text{при } x_t = 0. \end{cases} \quad (1)$$

Оценка вероятности появления единицы для «виртуального скользящего окна» $\hat{p}_{t+1} = \frac{s_t}{cW}$. Параметр алгоритма c и начальное значение s_0 выбирается согласно следующему утверждению.

Утверждение. Для того, чтобы минимальные и максимальные оценки вероятностей для «скользящего окна» и «виртуального скользящего окна» совпадали, необходимо и достаточно, чтобы параметр $c = W - 1 - \frac{2}{W}$ и перед кодированием символа x_1 последовательности x_1, x_2, \dots, x_n начальное значение $s_0 \in \{\frac{W}{2} - 1, \dots, cW - \frac{W}{2} + 1\}$. ■ При $W \gg 1$, можно считать, что $c = W - 1 - \frac{2}{W} \approx W$.

При использовании алгоритма «виртуального скользящего окна» каждому контексту может быть сопоставлена длина окна, учитывающая статистические свойства соответствующего двоичного источника. Для этого в диссертационной работе предложен следующий подход. Сначала кодируется некоторое множество тестовых видеопоследовательностей, на которых определяются наилучшие длины окон для каждого контекста. Затем полученные длины окон используются при кодировании других видеопоследовательностей.

Использование алгоритма «виртуального скользящего окна» позволяет уменьшить скорость кодирования на 0.5-1.5% при фиксированном уровне искажения по сравнению с табличными алгоритмами оценки вероятности, входящими в стандарт H.264/AVC и JPEG2000. При этом алгоритм «виртуального скользящего окна» может быть реализован как без использования таблиц, так и без операций умножения и деления (при $W = 2^i$, где i - неотрицательное целое число).

Выбор длины окна для каждого контекста осуществлялся из предположения, что наилучшая длина окна одинакова для двоичных источников, соответствующих одному контексту различных видеопоследовательностей. В общем случае это предположение не всегда выполняется и возникает задача разработки алгоритмов оценки вероятности, степень сжатия для которых близка к степени сжатия, достигаемой при заранее известной наилучшей длине окна для данной двоичной последовательности.

В конце раздела описано как для решения поставленной задачи может быть использован метод универсального кодирования с учетом функции цели, предложенный Ю.М. Штарьковым. Для выбранной функции цели доказывается утверждение, что достигаемая степень сжатия с ростом длины кодируемой двоичной последовательности приближается к степени сжатия для «виртуального скользящего окна» в случае заранее известной наилучшей длины окна. Предложенный подход позволяет получить дополнительное уменьшение скорости кодирования до 0.5% по отношению к алгоритму «виртуального скользящего окна».

Третий раздел посвящен задаче повышения эффективности передачи видеоданных по каналам связи с низкой пропускной способностью при помощи усовершенствования методов устранения временной избыточности источников видеoinформации, основанных на блоковой оценке движения.

Наиболее распространенный алгоритм блоковой оценки движения упрощенно может быть описан следующим образом. Для каждого блока S_n в текущем (кодируемом) кадре с номером n выполняется поиск блока S_k в предыдущем (базовом) кадре с номером k , соответствующий минимуму следующей целевой

функции

$$J(\mathbf{v}) = \sum_{(x,y) \in S_n} |s_n(x, y) - s_k^*(x + v_x, y + v_y)|, \quad (2)$$

где $\mathbf{v} = (v_x, v_y)$ - вектор движения, $|v_x| \leq r$ и $|v_y| \leq r$, r - радиус поиска, $s_n(x, y)$ - значение яркости пиксела с координатами x и y в кодируемом кадре, $s_k^*(x + v_x, y + v_y)$ - значение яркости пиксела с координатами $x + v_x$ и $y + v_y$ в базовом кадре. Затем, путем вычитания соответствующих значений яркостей и цветностей пикселов блоков S_n и S_k , формируется разностный блок, который кодируется вместе с вектором движения.

Целевая функция (2) не учитывает битовые затраты на векторы движения, доля которых может достигать 40-50% при высоких степенях сжатия. Поэтому для увеличения степени сжатия видеоданных при передаче по каналам связи с низкой пропускной способностью распространена оценка движения, в которой последовательно для каждого блока выполняется минимизация следующей целевой функции

$$\Omega(\mathbf{v}) = J(\mathbf{v}) + \lambda(q) \cdot r(\mathbf{v}), \quad (3)$$

где $r(\mathbf{v})$ - битовые затраты на вектор движения, $\lambda(q)$ - множитель, зависящий от шага квантования q . При последовательной минимизации (3) вектор движения, найденный для текущего блока, влияет на выбор вектора движения для следующего блока. Это связано с тем, что обычно кодируется не абсолютное значение компонент вектора v_x и v_y , а разность компонент вектора для текущего блока и вектора для предыдущего блока. Поэтому такая минимизация в общем случае не приводит к минимизации битовых затрат как на векторы движения, так и на разностные блоки для всего кадра.

В ряде работ для ускорения поиска векторов движения предлагается использовать «иерархическую оценку движения» при которой кодируемый и базовый кадры сначала уменьшаются в два раза по высоте и ширине (см. рис. 1). Полученные кадры снова уменьшаются и так далее. Данная процедура выполняется L раз, где L - параметр алгоритма. В результате формируются $L + 1$ уровней иерархии, которые нумеруются числами от 0 до L , причем исходные (кодируемый и базовый) кадры соответствуют нулевому уровню. Полученные на всех уровнях изображения разбиваются на блоки одинакового размера (например, 16×16 пикселов). В результате такого деления каждому блоку на уровне l соответствуют четыре блока на уровне $l - 1$ и так далее. Затем выполняется поиск вектора движения $\mathbf{v}(L, i, j)$ для блока с координатами (i, j) на уровне L . Найденный вектор используется при поиске четырех векторов на уровне $L - 1$ и так далее до нулевого уровня. Множество векторов движения $\{\mathbf{v}(0, i, j)\}$, соответствующее нулевому уровню иерархии, используется для формирования разностных блоков. При этом не используются все «промежуточные» векторы движения, найденные на остальных уровнях иерархии.

В диссертационной работе предлагается модификация алгоритма «иерархической оценки движения», согласно которой для формирования разностных блоков могут быть использованы как «промежуточные» векторы движения, так и векторы движения, найденные на нулевом уровне иерархии. При таком подходе вектор движения $\mathbf{v}(1, i, j)$, найденный на первом уровне иерархии, может быть использован при формировании 4^1 соответствующих

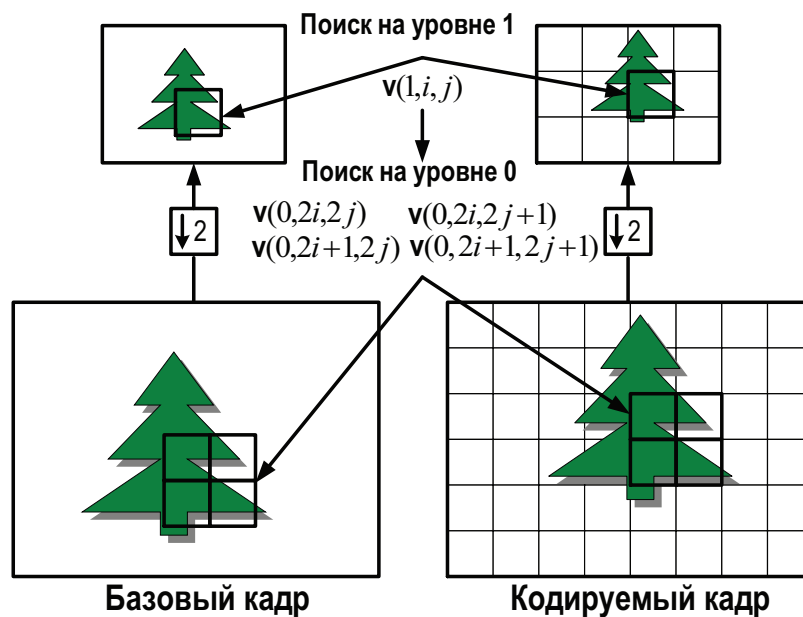


Рисунок 1 – Иерархическая оценка движения

блоков на нулевом уровне иерархии и так далее, вектор движения $\mathbf{v}(L, i, j)$ может быть использован при формировании 4^L соответствующих блоков на нулевом уровне иерархии. Совокупность из 4^L блоков на нулевом уровне иерархии, которая соответствует одному блоку на уровне иерархии L называется *сегментом*. Таким образом, для каждого сегмента кадра возникает множество вариантов разбиения на блоки переменного размера, причем каждому блоку соответствует один вектор движения (см. пример разбиения на рис. 2, а).

Для передачи вида разбиения сегмента используется иерархическое дерево разбиения (см. рис. 2, б) узел которого $t(l, i, j)$ принимает значение «0» в случае, если блок (i, j) на уровне l не разбивается и «1», если блок разбивается на четыре блока. Иерархическое дерево кодируется, начиная с корневой вершины. При этом если значение текущего узла равно «0», то дочерние узлы не кодируются.

Векторы движения предлагается кодировать в соответствии с деревом разбиения. Если узел дерева $t(l, i, j) = 1$, то кодируется разность между вектором $\mathbf{v}(l - 1, 2 \cdot i, 2 \cdot j)$, найденным на уровне $l - 1$, и удвоенным вектором $\mathbf{v}(l, i, j)$, найденным на уровне l .

В диссертационной работе предлагается процедура разбиения (определения значений в узлах иерархического дерева) сегмента кадра для которой битовые затраты на векторы движения, иерархическое дерево и разностные блоки для сегмента полностью определяются шагом квантования q . Доказано утверждение, согласно которому предлагаемая процедура минимизирует битовые затраты на векторы движения и иерархическое дерево при условии, что битовые затраты на разностные блоки минимальны для всех вариантов разбиения сегмента.

Использование предложенного алгоритма для стандарта MPEG-2 позволяет уменьшить скорость кодирования на 5-40% при фиксированном уровне искажения по сравнению с алгоритмами оценки движения, использующими целевые функции (2) и (3). При этом максимальный выигрыш по сжатию достигается в диапазоне низких скоростей кодирования (см. рис. 3).

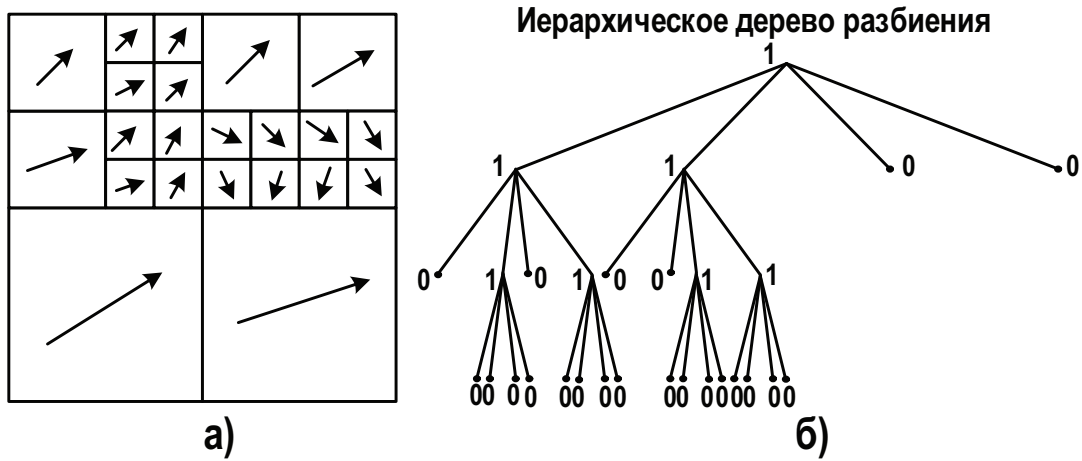


Рисунок 2 – Иерархическое разбиение сегмента

Описанные выше алгоритмы оценки движения имеют следующий существенный недостаток. Для этих алгоритмов заранее неизвестна точка на кривой скорость/искажение, которая получится в результате выбора векторов движения. Поэтому количество бит на кадр для этих алгоритмов, фактически, является случайной величиной. Данное свойство усложняет реализацию систем сжатия и передачи видеоинформации с жесткими ограничениями на задержку передачи данных. Однако предложенный в диссертационной работе модифицированный алгоритм «иерархической оценки движения» может быть обобщен для случая оценки движения с ограничением, при которой осуществляется совместный выбор векторов движения и шагов квантования при заданном ограничении на количество бит на кадр.

Обозначим за $r_i^{q_i}$ и $d_i^{q_i}$ битовые затраты и уровень искажения на сегмент i соответственно, при использовании шага квантования q_i . Обозначим за $\mathbf{q} = \{q_i\}$ вектор, компонента q_i которого означает, что для сегмента i используется шаг квантования q_i . Тогда суммарные битовые затраты на кадр составят $r(\mathbf{q}) = \sum_{i=1}^N r_i^{q_i}$, где N - количество сегментов в кадре. Аналогично, уровень искажения на кадр $d(\mathbf{q}) = \sum_{i=1}^N d_i^{q_i}$. Тогда при оценке движения с ограничением количества бит на кадр необходимо найти вектор \mathbf{q}^* так, чтобы

$$\begin{cases} d(\mathbf{q}^*) = \min_{\mathbf{q} \in \{\mathbf{q}\}} d(\mathbf{q}), \\ \text{при условии, что } r(\mathbf{q}^*) \leq r_{max}. \end{cases} \quad (4)$$

Решение данной задачи может быть получено при помощи метода «лагранжевых релаксаций», предложенного Г. Шустером и А. Катсаггелосом.

В конце раздела показывается как предложенный алгоритм оценки с ограничением может быть использован при управлении скоростью кодирования группы видеоисточников в системах цифрового телевизионного вещания. При традиционном подходе в таких системах пропускная способность канала связи равномерно делится между видеоисточниками. После этого при помощи алгоритма управления скоростью кодирования каждый видеоисточник кодируется в соответствии с выделенным ресурсом. Однако из-за статистических отличий видеоисточников такой подход может привести к существенной разнице уровней искажения между передаваемыми видеопоследовательностями.

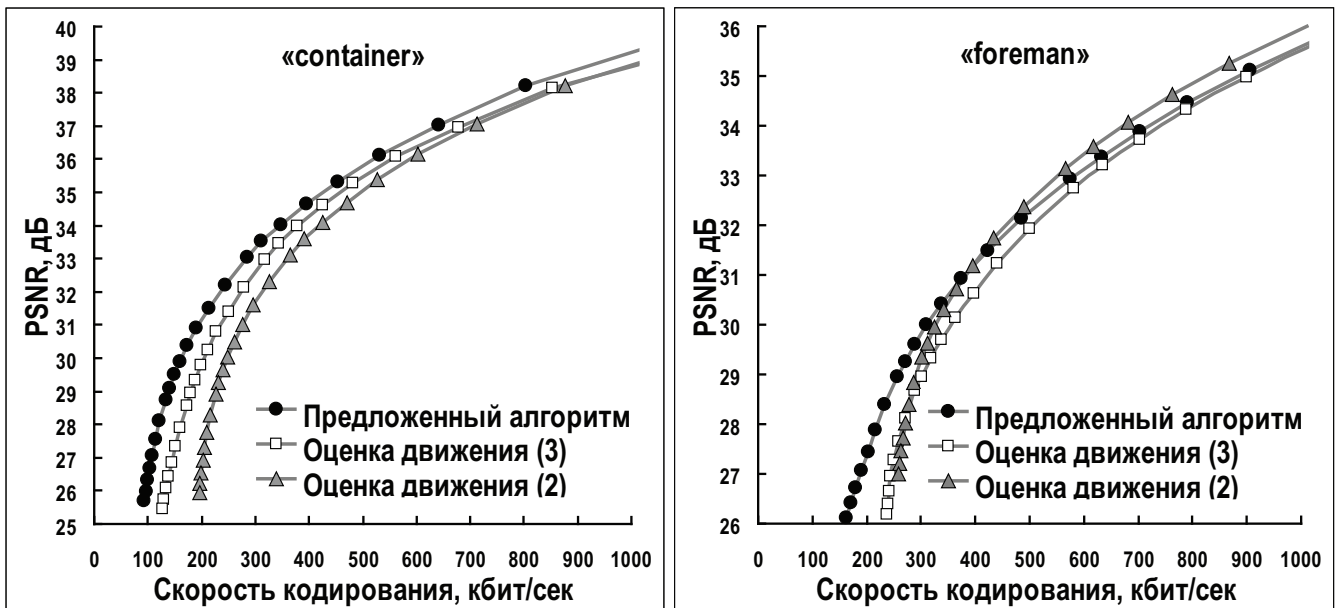


Рисунок 3 – Сравнительные результаты для MPEG-2

Для передачи видеоданных с одинаковым уровнем искажения предлагается общий ресурс канала связи распределять между видеоисточниками по минимаксному критерию искажения. Пусть имеются S видеоисточников. Обозначим за $d(s, \mathbf{q}_i)$ и $r(s, \mathbf{q}_i)$ уровень искажения и битовые затраты на кадр с номером i видеоисточника с номером s при использовании множества шагов квантования \mathbf{q}_i для сегментов в кадре. Поделим каждую видеопоследовательность на группы из M кадров. Группа кадров с номером j состоит из номеров кадров $G_j = \{j \cdot M, \dots, (j + 1) \cdot M - 1\}$.

Тогда для каждого кадра i , входящего в группу кадров с номером j источника s , необходимо выбрать множество шагов квантования \mathbf{q}_i^* так, чтобы

$$\begin{cases} \text{минимизировать } \max_{s \in S, i \in G_j} d(s, \mathbf{q}_i) \\ \sum_{s \in S} \sum_{i \in G_j} r(s, \mathbf{q}_i^*) \leq \frac{M}{f} \cdot c, \end{cases} \quad (5)$$

где f - число кадров в секунду для каждого видеоисточника, c - скорость канала связи.

Предлагаемое в диссертационной работе решение данной задачи основано на использовании модифицированного алгоритма «иерархической оценки движения» с ограничением.

В **четвертом** разделе рассматривается задача управления скоростью кодирования видеoinформации с целью передачи данных на небольшие расстояния по высокоскоростным каналам связи с малой задержкой передачи данных.

Высокая скорость канала связи позволяет осуществлять передачу видеопоследовательностей большого разрешения. При большом разрешении (например, 1920×1080) объем памяти как кодирующего, так и декодирующего устройств оказывается много меньше количества бит, необходимого для хранения одного кадра видеопоследовательности. В этих условиях на

вход кодирующего устройства подается небольшой сегмент входного кадра видеопоследовательности, который обрабатывается отдельно от остальных. При такой независимой обработке невозможна реализация методов устранения временной избыточности. Поэтому разрабатываемые алгоритмы могут достигать определенной степени сжатия только за счет устранения пространственной избыточности. В связи с этим рассматриваемые в разделе подходы реализованы на базе стандарта JPEG2000, который наиболее эффективен при таком ограничении.

Для формулирования требований к алгоритму управления скоростью кодирования в диссертационной работе вводится следующая модель системы передачи. Время передачи по каналу разделено на окна. Все окна имеют одинаковую длительность, равную времени передачи одного пакета. Окна пронумерованы целыми неотрицательными числами, окну с номером t соответствует интервал времени $[t, t + 1)$. Далее для краткости изложения окно с номером t называется окном t . Источник видеoinформации через одинаковые интервалы времени подает на вход кодера очередной сегмент видеокadra. Кодер работает в реальном масштабе времени, поэтому на момент окончания окна t в «сглаживающий» буфер передатчика помещается сжатый сегмент размером r_t бит. Из буфера передатчика данные передаются приемнику по каналу связи с постоянной скоростью c . В зависимости от количества бит в буфере передатчика алгоритм управления скоростью кодирования формирует параметры сжатия для следующего сегмента.

Количество бит в буфере передатчика $b^e(t)$ в окне t после помещения в буфер очередного сжатого сегмента объемом r_t и передачи по каналу связи изменяется следующим образом:

$$b^e(t) = \max\{0, b^e(t - 1) - c\} + r_t, \quad (6)$$

где c - скорость канала связи.

Данные на приемной стороне накапливаются некоторое время L , после чего начинается декодирование и воспроизведение. Суммарная задержка ΔT между моментом времени, когда на вход кодера поступил очередной сегмент и моментом времени, когда данный сегмент отобразился на экране приемной стороны, состоит из следующих компонент:

$$\Delta T = \Delta T_e + \Delta T_{be} + \Delta T_c + \Delta T_{bd} + \Delta T_d, \quad (7)$$

где ΔT_e и ΔT_d - время обработки сегмента кодером и декодером соответственно, ΔT_{be} - время ожидания передачи сегмента в буфере передатчика, ΔT_{de} - время ожидания воспроизведения сегмента в буфере передатчика, ΔT_c - время передачи сегмента по каналу связи. Известно, что если объем буфера на передающей и приемной сторонах $B_{max}^e = B_{max}^d = L \cdot c$, и устройство управления функционирует так, что количество бит в буфере передатчика $b^e(t) \leq B_{max}^e$ (буфер передатчика не переполняется), то задержка ΔT , вносимая при буферизации данных на приемной стороне, постоянна и равна задержке начальной буферизации L .

С учетом ограничения на задержку передачи данных требования к алгоритму управления скоростью кодирования в диссертационной работе сформулированы

в виде следующей *минимаксной оптимизационной задачи*. Для передаваемого в окне t сегмента необходимо выбрать такие параметры кодирования, чтобы

$$\begin{cases} \text{минимизировать } \max_t d_t \\ b(t) \leq B_{max}^e, \end{cases} \quad (8)$$

где d_t - уровень искажения для сегмента, передаваемого в окне t .

Решение данной задачи может быть найдено при помощи гипотетического алгоритма, который в диссертационной работе назван *алгоритмом последовательного поиска*. Обозначим за $\mathbf{D} = \{d_1, d_2, \dots\}$ множество всех возможных уровней искажения для всех сегментов, входящих в видеопоследовательность, упорядоченных по возрастанию. При последовательном поиске каждый сегмент сначала кодируется с искажением не выше d_1 , при этом согласно (6) вычисляется количество бит в буфере передатчика. Если произошло переполнение буфера, то вся процедура начинается сначала и все сегменты кодируются с ошибкой не выше d_2 и так далее, до тех пор пока при некотором $\tilde{d} \in \mathbf{D}$ не будет происходить переполнение буфера. Для последовательного поиска доказано следующее утверждение.

Утверждение. *Не существует последовательности параметров кодирования, не приводящих к переполнению буфера и имеющих максимальное значение уровня искажения меньше чем \tilde{d} , где \tilde{d} - максимальное значение искажения, полученное алгоритмом последовательного поиска. ■*

Алгоритм последовательного поиска не может быть реализован на практике, так как при передаче в реальном масштабе времени нет возможности начинать передачу заново после переполнения буфера. Поэтому в диссертационной работе предлагается эвристический алгоритм управления скоростью кодирования, который позволяет получить оценку значения \tilde{d} алгоритма последовательного поиска в процессе работы системы при ограничениях на память и задержку передачи данных (см. рис. 4). Обозначим эту оценку за $\hat{d}(t)$. Оценивать значение \tilde{d} предлагается следующим образом. До тех пор, пока количество бит в буфере не превысит некоторый порог V_H , все сегменты сжимаются с искажением не более $\hat{d}(t)$ (режим заполнения буфера). Превышение порога V_H означает, что поддержание искажения на уровне $\hat{d}(t)$ невозможно для данной пропускной способности канала связи. Поэтому все сегменты кодируются с максимальным, но приемлемым уровнем искажения d_{empty} до тех пор, пока не произойдет опустошение буфера (режим опустошения буфера). После этого оценка уровня искажения $\hat{d}(t)$ увеличивается на величину Δd , где Δd - параметр алгоритма, и система переходит в режим заполнения буфера. Для данного алгоритма доказано следующее утверждение.

Утверждение. *Пусть алгоритму последовательного поиска с объемом буфера \tilde{V} соответствует максимальный уровень искажения \tilde{d} . Тогда, при использовании предложенного алгоритма с порогом $V_H = \tilde{V}$ и $\hat{d}(0) < \tilde{d}$, для всех окон t выполняется неравенство $\hat{d}(t) \leq \tilde{d} + \Delta d$. ■*

Оптимизационная задача (8) предполагает, что статистические свойства отдельных сегментов изображения могут существенно отличаться друг от друга, но статистические свойства кадра в целом меняются незначительно.

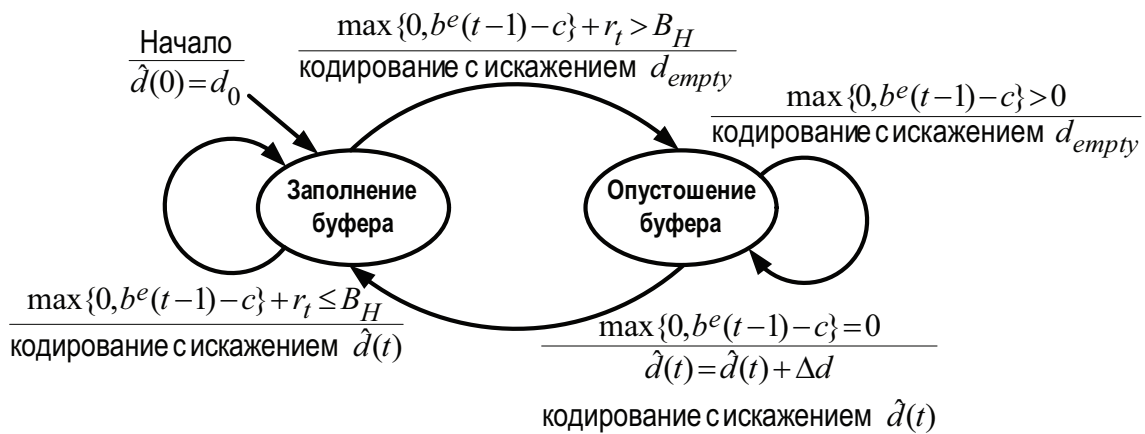


Рисунок 4 – Иллюстрация работы алгоритма управления

В общем случае, видеопоследовательность состоит из подпоследовательностей (или сцен) с различными статистическими свойствами. В связи с этим, в диссертационной работе вводится *расширенная минимаксная оптимизационная задача* при которой оптимизационная задача (8) применяется для каждой сцены. Затем аналогично вводится эвристический алгоритм управления скоростью кодирования с учетом смены сцены.

На видеопоследовательностях, содержащих компьютерную графику предложенный алгоритм управления скоростью кодирования позволяет повысить пиковое отношение сигнал/шум на 5-10дБ относительно исходного алгоритма управления, входящего в стандарт JPEG2000 (см. «test1» на рис. 5). При этом для видеопоследовательностей, содержащих только естественные (фотографические) изображения, предложенный алгоритм управления показывает схожие результаты с алгоритмом управления, входящим в стандарт JPEG2000 (см. «test2» на рис. 5).

В конце раздела рассматриваются задачи, связанные с тестированием алгоритмов управления скоростью кодирования. Традиционный способ тестирования ограничивается получением экспериментальных результатов на небольшой выборке тестовых видеопоследовательностей. Это связано с трудоемкостью такого тестирования из-за необходимости сжатия и восстановления тестовых видеоданных для получения соотношения между уровнем искажения и скоростью кодирования. Ограничение множества тестовых видеопоследовательностей препятствует как отладке программного обеспечения, на котором реализован алгоритм управления скоростью кодирования, так и получению характеристик алгоритма управления.

В диссертационной работе предлагается методика тестирования алгоритмов управления, в которой вместо сжатия и восстановления тестовых видеопоследовательностей используется модель источник/кодер видеоинформации. Модель описывает зависимость между уровнем искажения и битовыми затратами для очередного сегмента, при использовании заданного кодера видеоинформации. Такой подход позволяет существенно уменьшить время тестирования и время получения характеристик алгоритма управления.

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

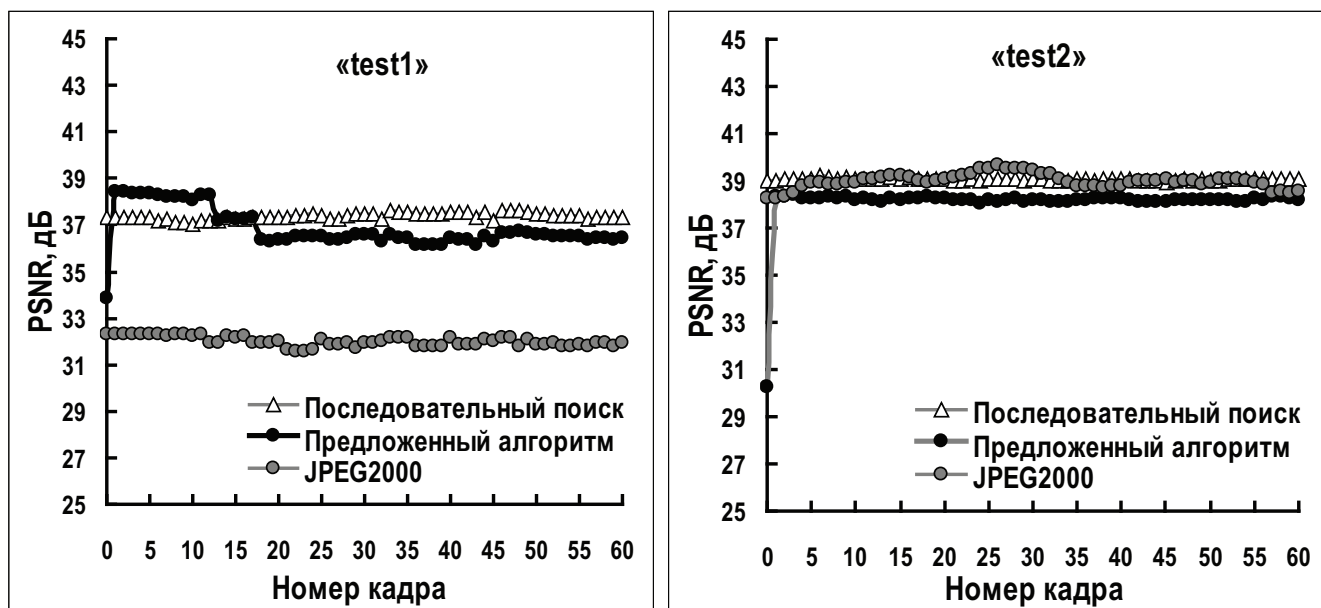


Рисунок 5 – Сравнение алгоритмов управления

В **приложении** приводится упрощенное описание стандартов сжатия видеоинформации, используемых в диссертационной работе, метод «лагранжевых релаксаций», практические результаты для разделов 3 и 4, а также два акта об использовании результатов диссертационной работы.

ЗАКЛЮЧЕНИЕ

В диссертационной работе получены следующие основные результаты:

1. Предложен алгоритм оценки вероятности при помощи «виртуального скользящего окна». Показано, что предложенный алгоритм позволяет увеличить степень сжатия достигаемого при адаптивном арифметическом кодировании относительно известных методов оценки вероятности.
2. Предложен модифицированный алгоритм «иерархической оценки движения». Показано, что предложенный алгоритм позволяет существенно снизить битовые затраты на векторы движения при высоких степенях сжатия по сравнению с известными алгоритмами оценки движения.
3. Рассмотрена задача управления скоростью кодирования для группы видеисточников по минимаксному критерию искажения. Предложено решение на основе модифицированного алгоритма «иерархической оценки движения».
4. Рассмотрена задача управления скоростью кодирования по минимаксному критерию искажения при ограничении на задержку передачи данных и объем памяти кодирующего и декодирующего устройств. Предложены алгоритмы управления, позволяющие повысить эффективность кодирования для видеопоследовательностей, содержащих фрагменты компьютерной графики.
5. Предложена методика тестирования алгоритмов управления скоростью кодирования при помощи модели источник/кодер видеоинформации.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Беляев Е.А., Тюрликов А.М. Оценка вероятности появления символа при адаптивном двоичном арифметическом кодировании в задачах сжатия видеоинформации // Цифровая обработка сигналов. 2007. №3. С.20-24.
2. Беляев Е.А., Тюрликов А.М. Управление скоростью и ошибкой кодирования в системе сжатия и передачи видеоинформации с ограничениями на память передающего и принимающего устройств // Компьютерная оптика. 2007. Т.31. №2. С.69-76.
3. Беляев Е.А., Тюрликов А.М., Уханова А.С. Адаптивное арифметическое кодирование в стандарте JPEG2000 // Информационно-управляющие системы. 2007. №6. С.28-33.
4. Беляев Е.А. Использование «виртуального скользящего окна» при адаптивном арифметическом кодировании // Научная сессия ГУАП. 2006. С.243-246.
5. Belyaev E., Gilmutdinov M., Turlikov A. Binary arithmetic coding system with adaptive probability estimation by «virtual sliding window» // 2006 IEEE 10th International Symposium on Consumer Electronics. 2006. P.194-198.
6. Беляев Е.А., Чуйков А.В. Модифицированный алгоритм иерархической оценки движения в задачах сжатия видеоинформации // Научная сессия ГУАП. 2007. С.56-60.
7. Belyaev E., Turlikov A., Ukhanova A. Rate-distortion control in wavelet-based video compression systems with memory restriction // XI International Symposium on Problems of Redundancy in Information and Control Systems. 2007. P.13-17.
8. Беляев Е.А. Определение момента смены сцены для улучшения алгоритма управления скоростью и ошибкой кодирования в системе сжатия и передачи видеоинформации при ограничении на память // Научная сессия ГУАП. 2008. С.95-98.
9. Беляев Е.А., Дубков К.С., Вычисление вероятности переполнения сглаживающего буфера передатчика для модели сжатия и передачи видеоинформации // Научная сессия ГУАП. 2008. С.99-101.
10. Беляев Е.А., Тюрликов А.М. Управление скоростью и ошибкой кодирования в системе сжатия и передачи видеоинформации в реальном масштабе времени при ограничении на память // 15-я Международная научно-техническая конференция «Проблемы передачи и обработки информации в сетях и системах телекоммуникаций». 2008. №1. С.72-73.
11. Belyaev E., Koski T., Paavola J., Turlikov A., Ukhanova A. Adaptive power saving on the receiver side in digital video broadcasting systems based on progressive video codecs // The 11th International Symposium on Wireless Personal Multimedia Communications. 2008. Lapland, Finland.
12. Belyaev E., Turlikov A., Ukhanova A. Rate-control algorithms testing by using video source model // The 15-th International Conference on Communications. 2008. St.Petersburg, Russia. P.1-5.