

На правах рукописи

Поляков Дмитрий Борисович

РАЗРАБОТКА МЕТОДОВ И УСТРОЙСТВ КОМПЕНСАЦИИ  
ПРОСТРАНСТВЕННО-ВРЕМЕННЫХ ИСКАЖЕНИЙ  
В СИСТЕМАХ ЦИФРОВОГО ТЕЛЕВИДЕНИЯ

Специальность 05.12.04 – Радиотехника,  
в том числе системы и устройства телевидения

АВТОРЕФЕРАТ  
диссертации на соискание ученой степени  
кандидата технических наук

Москва 2009

Работа выполнена на кафедре телевидения Государственного образовательного учреждения высшего профессионального образования Московский технический университет связи и информатики (ГОУ ВПО МТУСИ).

Научный руководитель: кандидат технических наук, доцент  
Игнатов Флегонт Михайлович

Официальные оппоненты: доктор технических наук, профессор  
Селиванов Арнольд Сергеевич

кандидат технических наук, доцент  
Смирнов Александр Витальевич

Ведущая организация: Всероссийский научно-исследовательский  
институт телевидения и радиовещания  
(ВНИИТР)

Защита состоится 19 марта 2009 г. в 15 часов на заседании совета по защите докторских и кандидатских диссертаций Д219.001.01 при ГОУ ВПО Московский технический университет связи и информатики по адресу: 111024, Москва, ул. Авиамоторная, д. 8а, ауд. А-455.

С диссертацией можно ознакомиться в библиотеке Московского технического университета связи и информатики.

Автореферат разослан « \_\_\_ » \_\_\_\_\_ 2009 г.

Ученый секретарь совета по защите  
докторских и кандидатских  
диссертаций,  
к.т.н., доцент

\_\_\_\_\_ Р.Ю. Иванюшкин

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы.** С широким распространением таких технологий, как цифровое телевидение, потоковое видео через Интернет и т.д., компрессия видеоданных, т.е. их обработка с целью сокращения объема информации, стала существенным компонентом широковещательного и развлекательного медиа. Успешное внедрение цифрового телевидения стало возможным, в частности, благодаря стандарту MPEG-2, разработанному более 15 лет назад и широко используемому в настоящее время, но морально устаревшему. На данный момент созданы более совершенные стандарты видеокодирования, например, H.264/AVC. На основании анализа рынка аппаратных разработок в области видеокодирования и направлений научных исследований можно сделать вывод, что он займет лидирующее положение среди форматов сжатия видео в ближайшее время. Однако применение более совершенных методов кодирования не позволяет в полной мере избежать появления характерных для видеокodeков искажений. Так как компрессия видеоданных – это всегда компромисс между коэффициентом сжатия и потерями в качестве, то такой процесс сопровождается возникновением в кодированном сигнале искажений, как пространственного, так и временного характера.

Специфика таких искажений различна для областей использования стандартов видеокодирования в широком диапазоне: от захвата и воспроизведения видео на мобильных устройствах с ограниченными вычислительными ресурсами до HDTV кинотеатров. Стоит отметить, что проблема компенсации и устранения искажений становится актуальной при кодировании с достаточно низкой скоростью битового потока. Особенную актуальность подобные исследования приобретают в связи с переходом России на цифровое телевидение.

Существенный вклад в разработку цифровых систем сжатия и обработки изображений, а также методов и аппаратуры измерения качества в цифровых телевизионных (ТВ) системах, внесли теоретические и практические работы отечественных ученых: С.И. Катаева, М.И. Кривошеева, В.Н. Безрукова, А.С. Селиванова, Б.П. Хромого, Ю.Б. Зубарева, В.П. Дворковича, а также исследования ряда зарубежных специалистов: Г. Салливана, Т. Вайгэнда, Дж. Вудса, А. Таураписа, А. Бовика и других.

Несмотря на растущий объем памяти носителей информации и увеличение пропускной способности каналов связи, повышение эффективности сжатия видеоматериала по-прежнему остается актуальной задачей, поскольку требования к качеству видео растут вместе с доступными вычислительными ресурсами. Задачи увеличения эффективности кодирования неразрывно связаны с задачами повышения качества изображения, поэтому проведение исследований и разработка алгоритмов компенсации и устранения искажений в кодированном сигнале телевизионных изображений является актуальным направлением исследований.

**Цель и задачи работы.** Целью настоящей диссертационной работы является разработка эффективных методов и устройств компенсации пространственных и временных искажений в системах цифрового телевидения, учитывающих специфику кодирования и преобразования сигналов, а также разработка и моделирование алгоритмов, реализующих предложенные методы. Для достижения поставленной цели в работе решаются следующие научно-практические задачи:

1. Анализ архитектуры современных стандартизированных и альтернативных видеокодеков с целью выявления характерных искажений в кодированном сигнале ТВ изображений.
2. Исследование и разработка алгоритмов компенсации пространственных искажений в системах сжатия статических и динамических изображений с поблочной обработкой изображения;
3. Исследование и разработка методов компенсации временных искажений в системах цифрового телевидения;
4. Исследование и разработка методов эффективного сжатия видеопоследовательностей с наличием локальных изменений сигнала яркости.

**Методы исследования.** При решении поставленных задач в работе использованы современные методы цифровой обработки изображений, статистической радиотехники, спектрального анализа и программирования.

**Личный вклад.** Все основные научные результаты, изложенные в диссертации, получены автором лично.

#### **Научная новизна работы.**

1. Получены аналитические выражения для частотных характеристик фильтров интерполяции в стандартах H.264/AVC, MPEG-4 Visual, AVS, на основе которых проведена оценка эффективности фильтрации.
2. Разработан алгоритм устранения блочных искажений на основе селективной частотной фильтрации.
3. Предложена модификация методики выбора режима внутрикадрового предсказания и режима кодирования макроблоков с целью уменьшения эффекта мерцания при использовании кодеков стандарта H.264/AVC.
4. Разработан алгоритм селективного взвешенного предсказания с целью повышения эффективности кодирования видеопоследовательностей с наличием локальных изменений сигнала яркости.

#### **Практическая ценность работы.**

1. Разработанный алгоритм устранения блочных искажений с селективной частотной фильтрацией встроен в программный кодек H.263+.
2. Использование алгоритма селективного взвешенного предсказания и модифицированной методики выбора режима внутрикадрового предсказания и режима кодирования макроблоков в кодеках стандарта H.264/AVC не требует изменения синтаксиса стандартного битового потока.

3. Разработаны структурные схемы для устройств селективного взвешенного предсказания и деблокинга на основе селективной частотной фильтрации.

**Реализация результатов работы.** Результаты диссертации использованы при разработке специальных устройств телевизионного контроля объектов во ФГУП МКБ «Электрон», а также внедрены в учебный процесс Московского технического университета связи и информатики в рамках дисциплин «Телевидение» и «Видеотехника». Реализация результатов работы подтверждена соответствующими актами.

**Апробация результатов работы.** Основные положения и результаты диссертационной работы докладывались и обсуждались на конференции профессорско-преподавательского, научного и инженерно-технического состава МТУСИ, 2006 г.; на Московских отраслевых научно-технических конференциях «Технологии информационного общества», 2007, 2008 гг.; на 5-й и 7-й Международных научно-технических конференциях «Фундаментальные проблемы радиоэлектронного и оптоэлектронного приборостроения» Intermatic, 2006, 2008 гг.; на 14-й Всероссийской и 16-й Международной научно-технических конференциях «Современное телевидение», 2006, 2008 гг.; на 63-й Научной сессии РНТОРЭС им. Попова, 2008 г.; на 5-й Международной научно-технической конференции «Молодые ученые – науке, технологиям и профессиональному образованию», 2008 г.

**Публикации.** По теме диссертации опубликовано 13 печатных работ, в том числе две статьи в ведущих рецензируемых журналах, входящих в перечень ВАК.

**Структура и объем работы.** Диссертационная работа состоит из введения, четырех глав, заключения, списка литературы и трех приложений, изложена на 193 страницах, иллюстрирована 86 рисунками, содержит 19 таблиц. Библиография включает 174 наименования.

**Основные положения, выносимые на защиту.**

1. Результаты оценки эффективности интерполяционной фильтрации в стандартах H.264/AVC, MPEG-4 Visual, AVS.
2. Алгоритм устранения блочных искажений на основе селективной частотной фильтрации позволяет повысить качество декодированных видеопоследовательностей.
3. Модифицированная методика выбора режима внутрикадрового предсказания и режима кодирования макроблоков уменьшает эффект мерцания при кодировании видеопоследовательностей по стандарту H.264/AVC.
4. Алгоритм селективного взвешенного предсказания позволяет увеличить эффективность кодирования по стандарту H.264/AVC видеопоследовательностей с наличием локальных изменений сигнала яркости.
5. Результаты экспериментального исследования разработанных алгоритмов.

## КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

**Во введении** обоснована актуальность темы диссертационной работы, указана цель работы, определены задачи и методы исследования. Сформулированы научная новизна, практическая значимость и положения, выносимые на защиту, а также кратко излагается содержание работы по главам.

**В первой главе** рассмотрена архитектура современных видеокодеков, проведен обзор перспективных направлений видеокодирования: многоакурсного видеокодирования и кодирования с генерацией масштабируемого битового потока. Проанализированы методики межкадровой и внутрикадровой обработки сигналов ТВ изображений в стандартных и альтернативных видеокодеках.

Рассмотрены архитектуры видеокодеков (Dirac, 3D SPIHT, MC-EZBC), основанных на вейвлет-преобразовании. Показано, что характерными для трехмерных вейвлет-кодеков временными искажениями при сжатии с большим коэффициентом компрессии, являются флуктуации объективных показателей качества во временных субполосах группы кадров, что вызывает неприятный для пользователя эффект мерцания. При очень низких скоростях кодирования становится заметна мешающая временная частота, связанная с накоплением ошибки в группе кадров.

Проведено профилирование кода эталонного программного обеспечения JM 13.2, отвечающего требованиям стандарта H.264/AVC. Установлено, что более 62% процессорного времени в кодере при использовании инструментов кодирования базового профиля H.264/AVC тратится на оценку движения. В декодере наиболее ресурсоемкими операциями являются интерполяционная (23,4%) и деблокинговая (22,5%) фильтрации.

Проведена оценка эффективности фильтров интерполяции в стандартах MPEG-4 Visual, H.264/AVC и AVS, применение которых необходимо для оценивания векторов движения с точностью до  $\frac{1}{2}$  элемента изображения. Импульсные характеристики данных фильтров  $h_{MPEG4}(k)$ ,  $h_{H264}(k)$ ,  $h_{AVS}(k)$ , где  $k$  – относительные координаты, определены в соответствующих стандартах как  $h_{MPEG4}(k) = [-8, 24, -48, 160, 160, -48, 24, -8]/256$ ,  $h_{AVS}(k) = [-1, 5, 5, -1]/8$ ,  $h_{H264}(k) = [1, -5, 20, 20, -5, 1]/32$ . Аппроксимированные кубическим сплайном импульсные характеристики построены на рис. 1(а). Для частотных характеристик фильтров получены аналитические выражения (1)-(3), по которым произведена оценка уровня подавления муаровых искажений (табл. 1).

$$H_{MPEG4}(\omega) = 1 + \frac{5}{4} \cos(\omega) - \frac{3}{8} \cos(3\omega) + \frac{3}{16} \cos(5\omega) - \frac{1}{16} \cos(7\omega), \quad (1)$$

$$H_{H264}(\omega) = 1 + \frac{5}{4} \cos(\omega) - \frac{5}{16} \cos(3\omega) + \frac{1}{16} \cos(5\omega), \quad (2)$$

$$H_{AVS}(\omega) = 1 + \frac{5}{4} \cos(\omega) - \frac{1}{4} \cos(3\omega), \quad (3)$$

где  $\omega$  – круговая пространственная частота,  $\omega \in [-\pi, \pi]$ .

Нормированные амплитудно-частотные характеристики фильтров интерполяции построены на рис. 1(б).

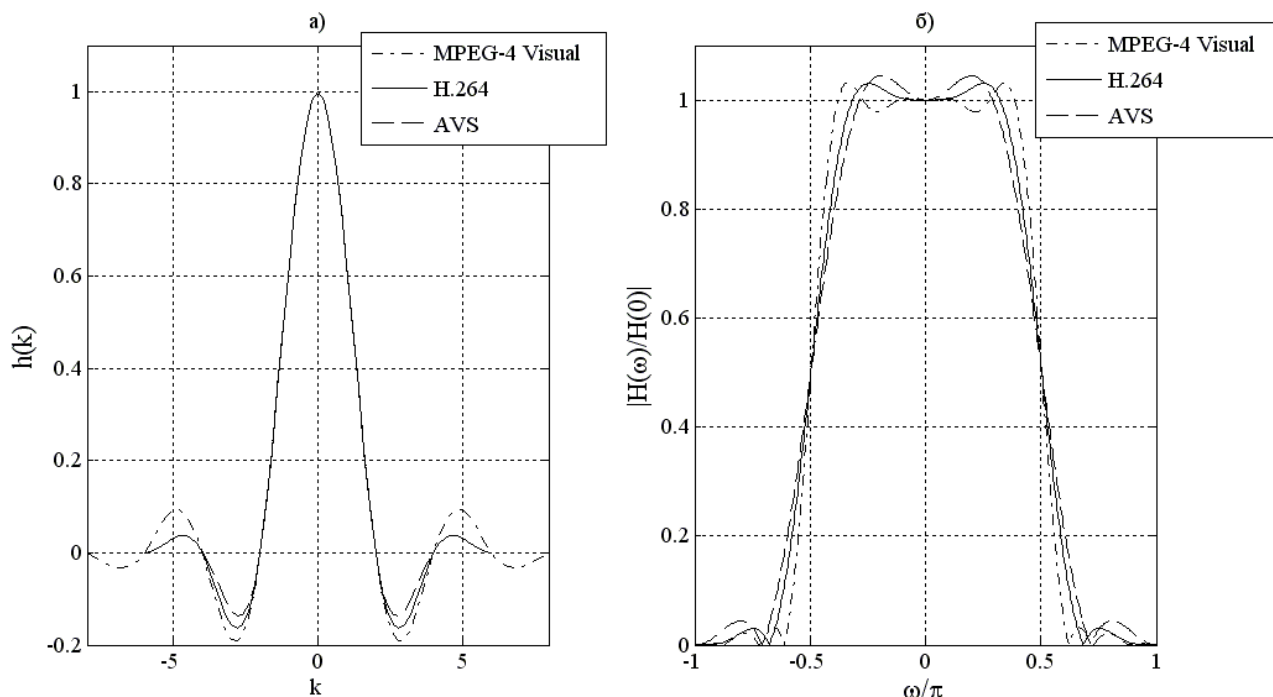


Рис. 1. Импульсные (а) и амплитудно-частотные (б) характеристики фильтров интерполяции

Согласно результатам исследования, наиболее качественный фильтр применяется в стандарте MPEG-4 Visual, в то время как фильтр стандарта H.264/AVC, обеспечивающий меньший уровень подавления муаровых искажений, имеет меньшую вычислительную сложность.

Таблица 1

Уровень подавления муаровых искажений фильтрами интерполяции, дБ

MPEG-4 Visual	H.264	AVS
17,8	15,9	14,9

Сделан вывод о том, что наиболее серьезным пространственным искажением в системах сжатия, основанных на поблочной обработке изображения, при кодировании с высоким коэффициентом компрессии, являются блочные искажения (блокинг). Также, при циклической вставке I-кадров, становятся заметны временные искажения в виде мерцания, одной из причин которых в кодеках стандарта H.264/AVC является наличие нескольких режимов внутрикадрового предсказания.

**Во второй главе** проведены исследования и разработка алгоритмов устранения пространственных и временных искажений.

Рассмотрены и классифицированы алгоритмы устранения блочных искажений (рис. 2).

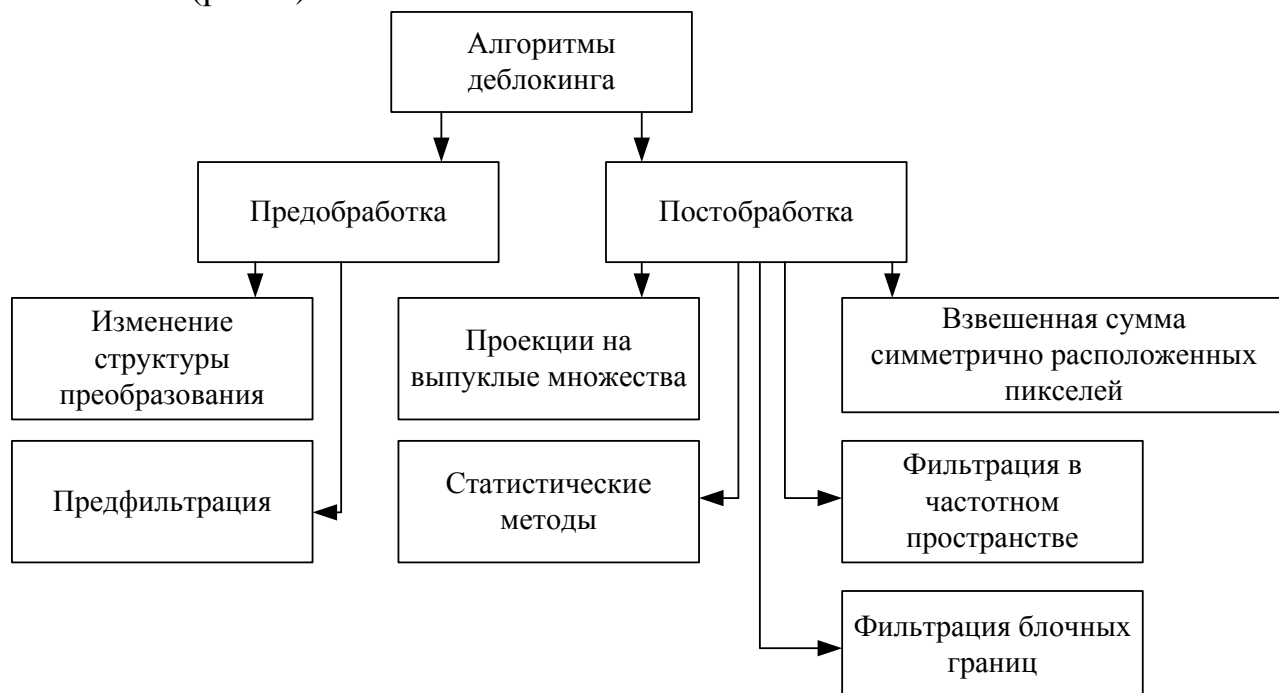


Рис. 2. Классификация алгоритмов устранения блочных искажений

Результаты анализа фильтров устранения блочных искажений позволяют сделать вывод, что достижение высокого качества восстановленного изображения отмечается при соблюдении следующих условий:

- вовлечении всех пикселей изображения в процесс фильтрации;
- осуществлении фильтрации в частотном пространстве.

Проведено сравнение эффективности работы фильтров устранения блочных искажений в стандартах H.263+, MPEG-4 Visual, H.264/AVC и VC-1 по критерию пикового отношения сигнал/шум (ПОСШ) и критерию структурного подобия (SSIM). Графики полученных зависимостей приведены в Приложении 1 к диссертационной работе. Измерена также вычислительная сложность алгоритмов с учетом количества инструкций, затрачиваемых на обработку одного кадра. По результатам исследования сделаны следующие выводы:

- наиболее эффективный фильтр – в стандарте H.264/AVC: он позволяет повысить качество изображения стандартной четкости в среднем до 0,46 дБ (до 0,0182 по критерию структурного подобия);
- наименее затратные по объему вычислений – алгоритмы деблокинга в стандартах H.263+ и VC-1;
- отсутствие фильтрации в петле кодирования следует признать недостатком архитектуры кодеков MPEG-4 Visual.

Предложена модификация алгоритма деблокинга, рассматривающего задачу устранения блочных искажений как задачу удаления шума из изображений. Такие алгоритмы перед обратным ДКП добавляют равномерный шум к коэффициентам преобразования, получая после обратного ДКП изображение, содержащее гауссовский шум. Любой из эффективных методов



удаления такого типа шума может использоваться на заключительном шаге алгоритма.

Существующие алгоритмы, предлагающие такой подход, классифицируют блоки декодированного изображения в реальной области путем подсчета их дисперсии и сравнения полученной величины с двумя заранее известными порогами. Если величина дисперсии блока ниже меньшего из порогов, то мощность добавляемого шума максимальна. При превышении дисперсией блока верхнего порога мощность шума минимальна, а в случае получения значения, лежащего между порогами, уровень добавляемого шума принимает некоторое среднее значение.

Недостатком такого подхода является необходимость обратного преобразования изображения в пространственную область для поблочного вычисления дисперсии изображения. Модификация алгоритма состоит в классификации блоков изображения в частотной области без выполнения обратного ДКП, что позволяет сократить общее время обработки. Предложено оценивать степень высокочастотного наполнения блока изображения путем подсчета количества ненулевых коэффициентов преобразования, в соответствии с которым адаптивно изменяется мощность вносимого шума.

Тестирование предложенного алгоритма проведено с использованием трех методов удаления шума из изображений: винеровской фильтрации, билатеральной фильтрации и фильтра нелокального усреднения. Применение таких методов обусловлено необходимостью предотвращения размывания текстурных областей, границ и контуров изображения. Результаты исследования приведены на рис. 3.

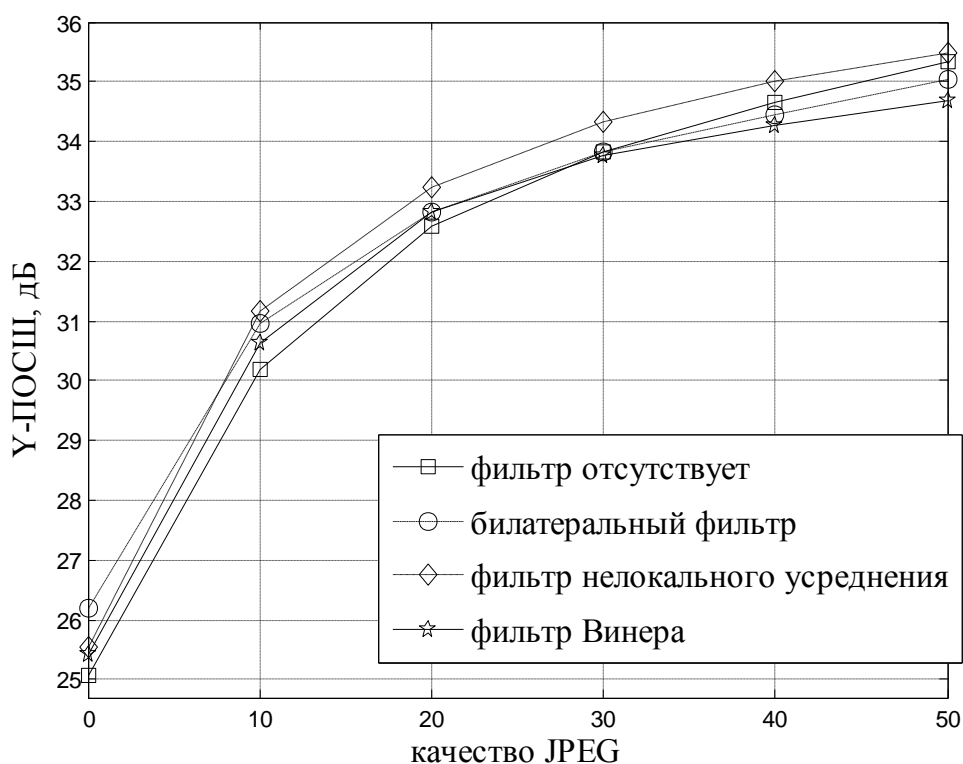


Рис. 3. Эффективность работы алгоритма деблокинга с инъекцией шума

По результатам исследования сделаны следующие выводы:

- 1) предложенный алгоритм обеспечивает увеличение объективного качества изображений до 1 дБ;
- 2) алгоритм имеет высокую вычислительную сложность и низкую или отрицательную эффективность работы при относительно высоком качестве изображения перед обработкой.

Полученные результаты говорят о том, что широкое внедрение алгоритмов устранения блочных искажений с инъекцией шума нецелесообразно.

Предложен алгоритм устранения блочных искажений на основе селективной частотной фильтрации, который удовлетворяет двум ранее заявленным критериям. Основой алгоритма является обработка изображения перекрывающимися блоками  $16 \times 16$ , которая производится согласно следующей предложенной методике.

Пусть  $A$  – область изображения  $16 \times 16$  и  $B_i$ , где  $i=1,2,3,4$ , – неперекрывающиеся блоки  $8 \times 8$ , составляющие участок изображения  $A$ . Пусть  $X$  – матрица частотных коэффициентов, полученных после выполнения ДКП над  $A$ , и  $Y_i$  – матрицы частотных коэффициентов, полученных после выполнения ДКП над  $B_i$ . Экспериментально установлено, что функция взаимной корреляции матрицы  $X$  и какой-либо матрицы  $Y_i$ , увеличенной до размеров  $X$  с помощью, например, билинейной интерполяции, имеет ярковыраженный пик при нулевых смещениях (рис. 4).

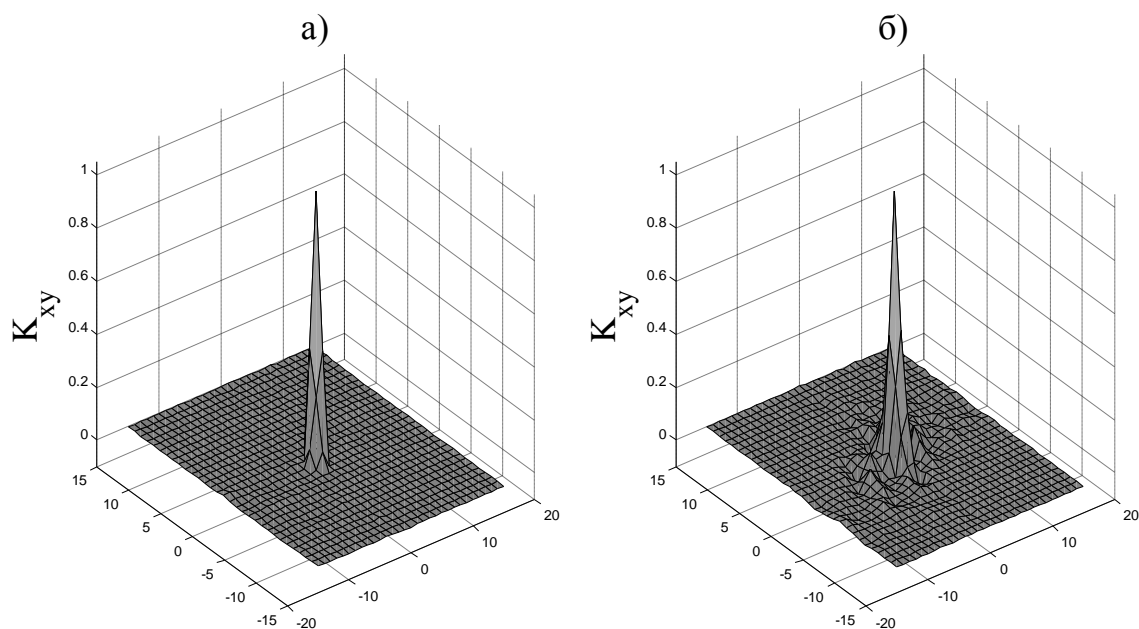


Рис. 4. Типичный вид функции кросс-корреляции между матрицами коэффициентов ДКП для фонового (а) и граничного (б) участков натуральных изображений

В связи с этим возможно проведение реконструкции элементов матрицы  $X$  по элементам  $Y_i$ . Однако с учетом аспектов вычислительной сложности и

точности предсказания с помощью билинейной интерполяции, предложено не реконструировать значения элементов  $x(i, j)$  матрицы  $\mathbf{X}$ , где  $i \in \{0, 1, \dots, 15\}$ ,  $j \in \{0, 1, \dots, 15\}$  - координаты элемента, а производить их фильтрацию по (4).

$$x(i, j) = \begin{cases} 0, & \forall m \in \{i-1, i, i+1\}, n \in \{j-1, j, j+1\} : z(m, n) = 0 \\ x(i, j), & \exists m \in \{i-1, i, i+1\}, n \in \{j-1, j, j+1\} : z(m, n) \neq 0, \end{cases} \quad (4)$$

где  $z(m, n)$  - элементы вспомогательной матрицы  $\mathbf{Z}$  размера  $16 \times 16$ , всем элементам которой изначально присваиваются нулевые значения, а элементы  $z(2k, 2l)$ , где  $k \in \{0, 1, \dots, 7\}$ ,  $l \in \{0, 1, \dots, 7\}$ , формируются согласно (5):

$$z(2k, 2l) = \begin{cases} 0, & y_1(k, l) = 0 \cap y_2(k, l) = 0 \cap y_3(k, l) = 0 \cap y_4(k, l) = 0 \\ 1, & y_1(k, l) \neq 0 \cup y_2(k, l) \neq 0 \cup y_3(k, l) \neq 0 \cup y_4(k, l) \neq 0, \end{cases} \quad (5)$$

где  $y_i(k, l)$  - элемент матрицы  $\mathbf{Y}_i$ .

Описанная выше процедура позволяет эффективно устранять блочные искажения в пределах блока обработки  $16 \times 16$ . Для перекрытия всех границ, остающихся необработанными на изображении после первого прохода алгоритма, потребуется дополнительно 3 раза обработать все изображение со смещенным начальным положением блока обработки  $16 \times 16$ . В каждом из дополнительных проходов обработки верхний левый угол блока обработки будет иметь координаты  $(0, 8)$ ,  $(8, 0)$  и  $(8, 8)$  соответственно.

Разработка весовой функции для взвешенного суммирования перекрывающихся блоков  $16 \times 16$  проводилась с привлечением экспериментального исследования и теоретического моделирования (рис. 5(а)) ошибки квантования в пространственной области при разных коэффициентах сжатия. Основываясь на полученных результатах, синтезирована весовая функция  $S(i, j)$  (рис. 5(б)), учитывающая зависимость ошибки квантования в пространственной области от координат в пределах блока  $16 \times 16$ .

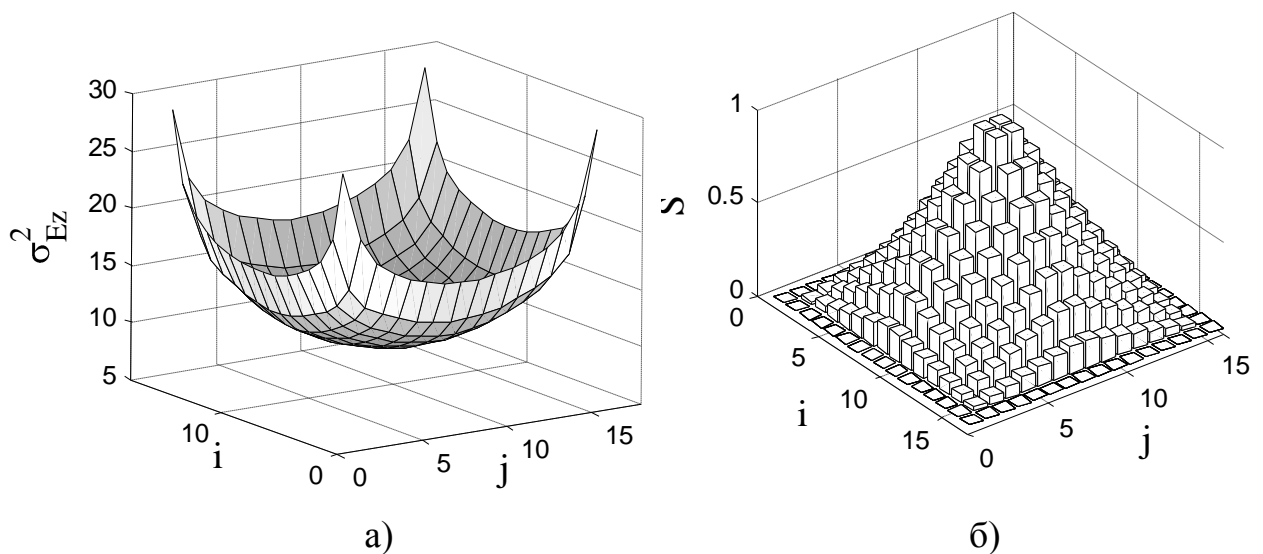


Рис. 5. Дисперсия ошибки квантования в пространственной области (а) и весовая функция (б)

Во второй главе также определены причины возникновения эффекта мерцания в кодеках стандарта H.264/AVC: присутствие пространственно-временного шума на видеопоследовательностях и наличие нескольких режимов внутрикадрового предсказания. Предложена модификация методики выбора режима внутрикадрового предсказания и режима кодирования макроблоков в стандарте H.264/AVC с целью уменьшения эффекта мерцания.

Как известно, в стандарте H.264/AVC используются 9 режимов внутрикадрового предсказания для блоков 4x4 и 4 режима для блоков 16x16. Лучший режим внутрикадрового предсказания  $I_p$  из  $p$  возможных для каждого из 16-ти блоков 4x4  $B_{4 \times 4}$ , составляющих текущий макроблок  $M$ , определяется по минимальному значению функции Лагранжа, вычисляемой согласно (6).

$$J(I_p, B_{4 \times 4}) = D(I_p, B_{4 \times 4}) + \lambda_{MODE} R(I_p, B_{4 \times 4}), \quad (6)$$

где  $D(I_p, B_{4 \times 4})$  - метрика искажений, определяемая как сумма квадратов разниц между значениями элементов исходного и декодированного блока, предсказанного в режиме  $I_p$ ;  $R(I_p, B_{4 \times 4})$  - количество бит, требуемых для кодирования блока  $B_{4 \times 4}$ , предсказанного в режиме  $I_p$ . Множитель  $\lambda_{MODE}$  определяется с помощью зависимости  $\lambda_{MODE}(QP) = 0.85 \times 2^{(QP-12)/3}$ , где  $QP$  - параметр квантования.

Лучший режим внутрикадрового предсказания для блока  $B_{16 \times 16}$ , составляющего текущий макроблок  $M$ , находится с помощью определения минимальной метрики искажений:

$$J(I_p, B_{16 \times 16}) = D(I_p, B_{16 \times 16}). \quad (7)$$

Решение о выборе режима кодирования макроблока I-кадра (Intra\_4x4 или Intra\_16x16) принимается также по минимальному значению функции Лагранжа:

$$J(I_p, M) = D(I_p, M) + \lambda_{MODE} R(I_p, M). \quad (8)$$

Установлено, что наличие нескольких режимов внутрикадрового предсказания является одной из причин появления эффекта мерцания в последовательностях, сжатых только с использованием внутрикадровых методов кодирования или с циклической вставкой I-кадров. Количественной оценкой этого эффекта служит индекс мерцания, определяемый для каждого кадра:

$$\hat{E}I = \frac{1}{N} \sum_{(i,j) \in \Omega_M} \max(0, \hat{d}_t(i,j) - d_t(i,j)), \quad (9)$$

где

$$\Omega_M = \{M_k \mid [\frac{1}{256} \sum_{(i,j) \in M_k} d_t(i,j)] < \varepsilon, k=1,2,\dots,N\}, \quad (10)$$

$$d_t(i,j) = |f_t(i,j) - f_{t-1}(i,j)|, \quad (11)$$

$$\hat{d}_t(i,j) = |\hat{f}_t(i,j) - \hat{f}_{t-1}(i,j)|, \quad (12)$$

$N$  - количество макроблоков в изображении, удовлетворяющих предъявленному в (10) критерию,  $M_k$  -  $k$ -тый макроблок,  $f_i$  и  $\hat{f}_i$  - исходный и декодированный кадры соответственно в момент времени  $t$ ,  $(i, j)$  - координаты элемента изображения. Порог  $\varepsilon$  определен экспериментально и равен 10.

Предлагаемая модификация методики принятия решений заключается в добавлении еще одного слагаемого к правой части функции Лагранжа, которое учитывает оценку величины эффекта мерцания для данного блока в текущем режиме. Таким образом, выражения (6) и (8) примут вид:

$$J(I_p, B_{4 \times 4}) = D(I_p, B_{4 \times 4}) + \lambda_{MODE} R(I_p, B_{4 \times 4}) + a \cdot \dot{E} \dot{I} (I_p, B_{4 \times 4}), \quad (13)$$

$$J(I_p, M) = D(I_p, M) + \lambda_{MODE} R(I_p, M) + b \cdot \dot{E} \dot{I} (I_p, \dot{I}), \quad (14)$$

где  $a$  и  $b$  – масштабирующие множители, определенные экспериментально. В проведенных экспериментах  $a=10$  и  $b=15$ .

**В третьей главе** исследован алгоритм взвешенного предсказания (ВП) в стандарте H.264/AVC, применение которого позволяет снизить вероятность определения ложного движения при наличии градационного перехода между кадрами видеопоследовательности.

ВП предусматривает возможность явного или неявного контроля относительного вклада опорных кадров в процесс компенсации движения. ВП оценивает яркость текущего пикселя мультипликативным весовым фактором  $w_{0C}$  и аддитивной величиной  $o_{0C}$  (15). При осуществлении межкадрового предсказания по списку опорных изображений  $L_0$ , ВП производится согласно (15):

$$predPart_C[x, y] = Clip1_C\{((predPartL0_C[x, y] \cdot w_{0C} + 2^{\log WD - 1}) \gg \log WD) + o_{0C}\}, \quad (15)$$

где  $predPartL0_C[x, y]$  – пиксель опорного изображения,  $predPart_C[x, y]$  – результат взвешенного предсказания, индекс  $C$  может принимать значения L, Cb, Cr (для яркостного и цветоразностных сигналов соответственно), функция  $Clip1_C$  ограничивает значения пикселей в соответствующем динамическом диапазоне, величина  $\log WD$  фактически играет роль квантователя (ее наличие продиктовано необходимостью иметь вес  $w_{0C}$  в целочисленной форме).

Экспериментально показано, что в случае глобального изменения яркости, т.е. при наличии градационного перехода между смежными кадрами, ВП позволяет закодировать текущее изображение, используя меньшее количество бит без потери в качестве. Однако при локальных изменениях яркости применение ВП дает неудовлетворительные результаты. В ходе экспериментов установлено, что при таких изменениях для кодирования кадра может потребоваться больше бит, чем при кодировании без ВП. Это является следствием того, что для каждого изображения возможно иметь только один набор параметров ВП.

Предложен алгоритм селективного ВП, учитывающий локальные изменения яркости и не требующий для своей реализации изменения синтаксиса битового потока в стандарте H.264/AVC.

Пусть  $P_{\delta \hat{a} \hat{e} \hat{o} \hat{u}}$  и  $P_{опорн}$  – текущее и опорное изображения соответственно, и  $(x, y)$  – координаты элемента изображения. Сначала производится оценка наличия градационного перехода между  $P_{\delta \hat{a} \hat{e} \hat{o} \hat{u}}$  и  $P_{\hat{i} \hat{v} \hat{i} \hat{\delta} \hat{r}}$ : определяется межкадровая разность (16), находится дисперсия полученной разности (17).

$$d(x, y) = P_{\delta \hat{a} \hat{e} \hat{o} \hat{u}}(x, y) - P_{\hat{i} \hat{v} \hat{i} \hat{\delta} \hat{r}}(x, y), \quad (16)$$

$$\sigma^2 = \frac{1}{N} \sum_{(x,y)} [d(x, y) - \bar{d}]^2, \quad (17)$$

где

$$\bar{d} = \frac{1}{N} \sum_{(x,y)} d(x, y), \quad (18)$$

$N$  – количество элементов изображения.

После определения  $\sigma^2$  проводится сравнение отношения  $\sigma^2 / N$  с пороговой величиной  $T1$ , превышение которой означает наличие сильного движения или смены сцены. В этом случае выполнение алгоритма прерывается, т.к. в обоих случаях его применение неэффективно. Порог определен экспериментально и равен 0,025. При  $\sigma^2 = 0$  алгоритм также заканчивает работу. При выполнении условия  $(\sigma^2 / N) < T1$  определяется критерий  $I$  (19) нормировкой дисперсии разности среднеквадратичной величиной этой разности для того, чтобы избежать определения ложных градационных переходов.

$$I = \frac{\sigma^2}{(E/N)}, \quad (19)$$

где

$$E = \sum_{(x,y)} [d(x, y)]^2. \quad (20)$$

Далее величина  $I$  сравнивается с порогом  $T2$ , при этом, если  $I < T2$ , то фиксируется наличие градационного перехода. Порог  $T2$  определен экспериментально и равен 0,92. При  $I \geq T2$  алгоритм заканчивает работу.

В качестве наименьшей единицы сегмента изображения принят макроблок, поскольку в его заголовке передается информация об индексе опорного изображения, по которому производится межкадровое предсказание.

Если величина  $D_M = \frac{1}{256} \sum_{(x,y) \in M} (P_{M_{текущ}}(x, y) - P_{M_{опорн}}(x, y))^2$ , определяемая для

каждого макроблока  $M$ , превышает значение отношения  $E/N$ , то этот макроблок относится к сегменту с высокой степенью относительных изменений, если нет – к сегменту с низкой степенью изменений. После вычисления величины  $D_M$  для всех макроблоков производится анализ окрестности каждого из макроблоков, отнесенных к первому сегменту. При наличии менее трех макроблоков из первого сегмента, смежных с текущим, маркировка принадлежности к первому сегменту с текущего макроблока снимается. Эта операция необходима для отбрасывания ложно отнесенных к

первому сегменту макроблоков, например, из-за наличия сильного движения между кадрами.

Определение параметров взвешенного предсказания предлагается производить по формулам (21) и (22).

$$w_{0L} = \text{floor}\left(\frac{P1_{N_k}}{P2_{N_k}} \cdot 2^{\log WD} + 0.5\right), \quad (21)$$

$$o_{0L} = \text{floor}\left(P1_{N_k} - \frac{w_{0L} \cdot P2_{N_k}}{2^{\log WD}} + 0.5\right), \quad (22)$$

где

$$P1_{N_k} = \sum_{(x,y) \in N_k} |P_{\text{текущи}_{N_k}}(x,y) - \bar{P}_{\text{текущи}_{N_k}}|, \quad (23)$$

$$P2_{N_k} = \sum_{(x,y) \in N_k} |P_{\text{опорн}_{N_k}}(x,y) - \bar{P}_{\text{опорн}_{N_k}}|, \quad (24)$$

$N_k$  –  $k$ -тый сегмент изображения,  $\bar{P}_{\text{текущи}_{N_k}}$  и  $\bar{P}_{\text{опорн}_{N_k}}$  – средние значения яркостей сегмента  $N_k$  в текущем и опорном изображениях соответственно,  $\text{floor}()$  – функция округления до ближайшего целого в меньшую сторону.

Обоснована возможность внедрения предложенного алгоритма селективного взвешенного предсказания в кодеки стандарта H.264/AVC без изменения синтаксиса битового потока, благодаря предусмотренному в стандарте процессу переупорядочивания списков опорных изображений. Этот процесс имеет возможность помещать в список изображения, которые не были в изначальном списке (при инициализации слайса). Таким образом, данный процесс позволяет не только переупорядочивать индексы в списках, но и модифицировать их так, чтобы несколько индексов фактически указывали на одно опорное изображение. При наличии нескольких копий одного опорного изображения (количество копий соответствует количеству сегментов, на которые разбивается изображение) с каждым из их индексов связывается собственный набор параметров ВП. Таким образом, никакого изменения синтаксиса битового потока не требуется, и все процедуры на приемной стороне производятся в соответствии со стандартом, а значит любой декодер, соответствующий стандарту H.264/AVC, способен декодировать битовый поток, созданный с использованием разработанного алгоритма селективного ВП.

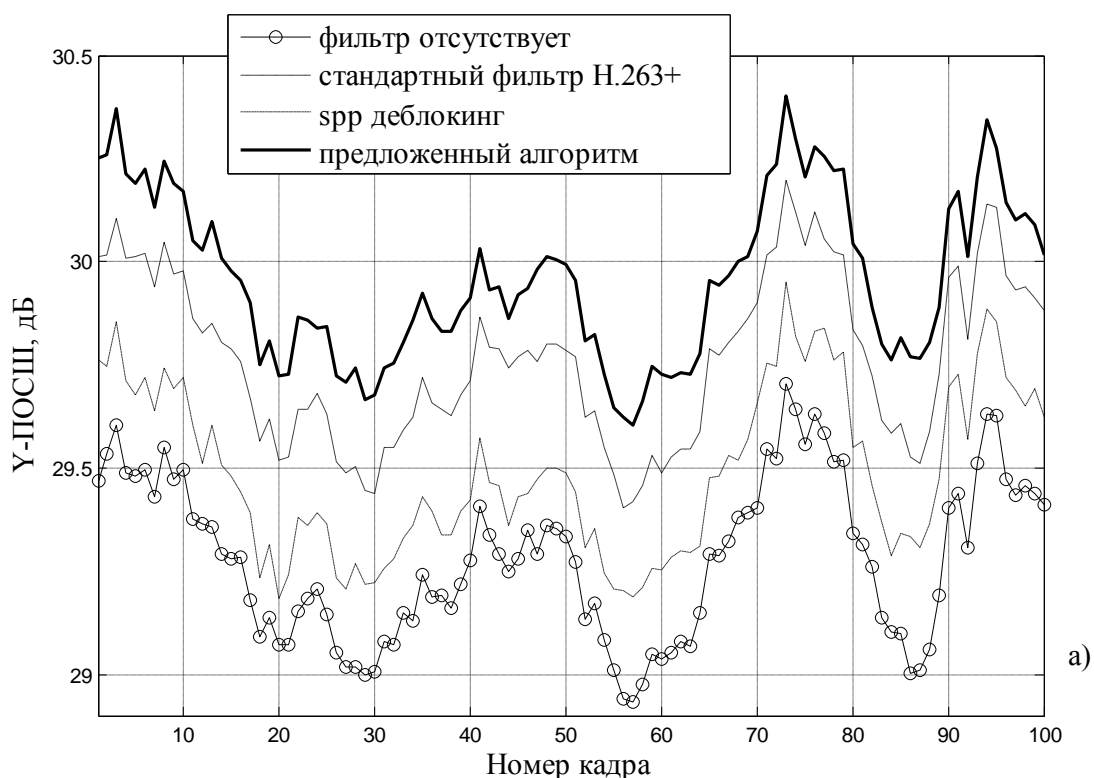
**В четвёртой главе** разработаны структурные схемы устройств, реализующих алгоритм деблокинга и алгоритм селективного взвешенного предсказания. Приведены результаты экспериментального исследования разработанных алгоритмов.

Алгоритм деблокинга на основе селективной частотной фильтрации встроен в программный кодек H.263+ (исходный код на языке Си для алгоритма приведен в Приложении 2 к диссертационной работе). С учетом

требования минимально возможной вычислительной и структурной сложности при программно-аппаратной реализации, в качестве ДКП 16x16 предложено использовать целочисленную аппроксимацию преобразования на основе факторизации Леффлера, позволяющей вычислять ДКП блока 16x16 за 13056 операций сдвига и 27136 операций сложения. Применение целочисленной аппроксимации ДКП позволяет избежать операций с плавающей точкой и упрощает практическую реализацию алгоритма на цифровых сигнальных процессорах.

Проведено сравнение разработанного алгоритма устранения блочных искажений со штатным фильтром деблокинга H.263+ и spr деблокингом, входящим в состав библиотеки DirectShow фильтров ffdshow. Визуализация объективных метрик качества (ПОСШ и SSIM) для 100 кадров видеопоследовательности «Foreman» приведена на рис. 6 (а) и (б) соответственно. Разработанный алгоритм устранения блочных искажений позволяет повысить качество изображений до 1 дБ по критерию ПОСШ и до 0,03 по критерию структурного подобия. Встроенными средствами языка Си оценено время обработки: для видеопоследовательности формата CIF (размер кадра 352x288 пикселей) задержка составляет около 10 мс на кадр.

Экспериментальные исследования модифицированного метода выбора режима внутрикадрового предсказания и режима кодирования макроблоков проводились с использованием кодека JM версии 14.0, отвечающего требованиям стандарта H.264/AVC. На рис. 7 представлены результаты работы предложенного метода при кодировании видеопоследовательности «Waterfall» с циклической вставкой I-кадров (25 кадров на группу) в формате IPP...P. В среднем уменьшение эффекта мерцания составило 24%. При этом потери в объективном качестве (по метрике ПОСШ) находятся в пределах 0.2 дБ.





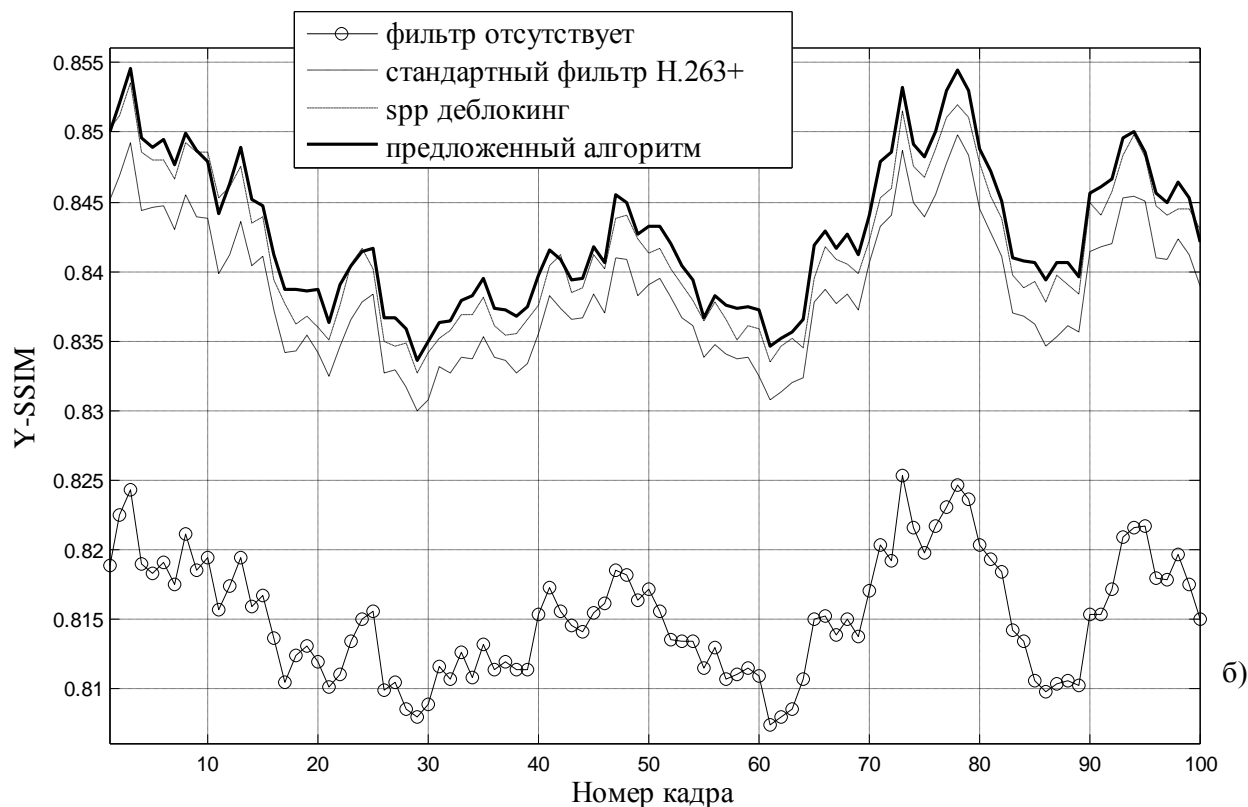


Рис. 6. Результаты сравнения алгоритмов деблокинга по критериям ПОСШ (а) и SSIM (б)

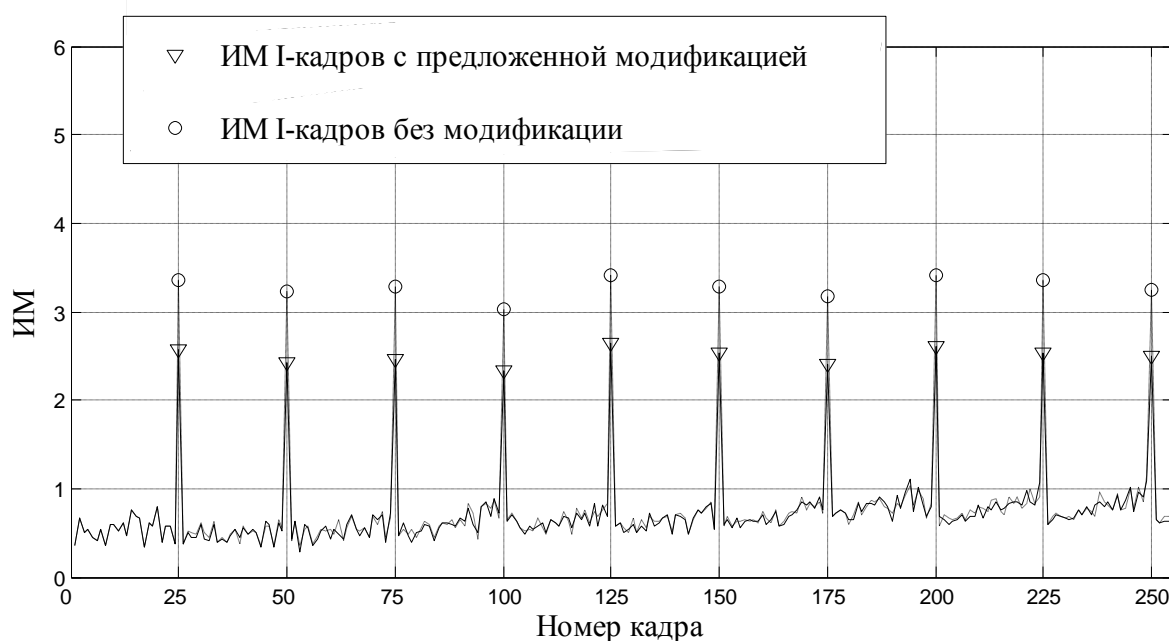


Рис. 7. Уменьшение эффекта мерцания при использовании модифицированной методики выбора режимов внутрикадрового предсказания и режима кодирования макроблоков

Экспериментальные исследования алгоритма селективного взвешенного предсказания проводились также с использованием кодека JM 14.0. В таблице 2 приведены результаты различных вариантов кодирования 29-го кадра последовательности «Stefan» со вспышкой на переднем плане. Кодирование

производилось с параметром квантования, равным 30. При использовании предложенного алгоритма потребовалось на 6% меньше бит для кодирования кадра.

Таблица 2

Эффективность работы алгоритма селективного взвешенного предсказания

Варианты кодирования	Количество бит на изображение	ПОСШ, дБ
без ВП	59576	34,63
алгоритм ВП по умолчанию	61992	34,67
предложенный алгоритм	56440	34,64

**В заключении** изложены наиболее значимые результаты, полученные в диссертационной работе, а именно:

1. Получены аналитические выражения для частотных характеристик фильтров интерполяции в стандартах H.264/AVC, MPEG-4 Visual, AVS, на основе которых проведена оценка эффективности фильтрации.
2. Установлено, что основными искажениями вейвлет-видеокодексов при кодировании с высоким коэффициентом компрессии является размытие изображения и колебание показателей объективного качества изображения в пределах группы кадров. Основными пространственными искажениями систем сжатия на основе ДКП являются блочные искажения и размытие изображения в пределах блока обработки. Характерным для систем сжатия на основе ДКП временным искажением является эффект мерцания, заметный для наблюдателя при достаточно низкой скорости кодирования видеопоследовательности с использованием только I-кадров или с циклической вставкой I-кадров.
3. Сравнение эффективности работы фильтров устранения блочных искажений в стандартах H.263+, MPEG-4 Visual, H.264/AVC и VC-1 показало, что наиболее эффективный фильтр среди сравниваемых – в стандарте H.264/AVC.
4. Установлено, что алгоритмы устранения блочных искажений с добавлением шума имеют высокую вычислительную сложность, низкое качество работы и малоприспособны для практической реализации.
5. Разработанный алгоритм устранения блочных искажений на основе селективной частотной фильтрации позволяет повысить качество изображений до 1 дБ по критерию пикового отношения сигнал/шум и до 0,03 по критерию структурного подобия.
6. Предложенная модификация методики выбора режима внутрикадрового предсказания и режима кодирования макроблоков в стандарте H.264/AVC позволяет снизить проявление эффекта мерцания в среднем на 24%.
7. Предложенный алгоритм селективного взвешенного предсказания позволяет сократить объем передаваемой информации при кодировании

кадра с наличием локальных изменений яркости в среднем на 6%. Обоснована возможность внедрения предложенного алгоритма в кодеки стандарта H.264/AVC без изменения синтаксиса стандартного битового потока.

### СПИСОК ПУБЛИКАЦИЙ

1. Поляков Д.Б. Сравнение алгоритмов деблокинга в стандартах видеокompрессии H.263, MPEG-4 ASP, MPEG-4 AVC, VC-1// Инфокоммуникационные технологии. – 2008. – Т.6. – № 3. – С. 31-35.
2. Поляков Д.Б. Повышение эффективности сжатия видео с использованием селективного взвешенного предсказания // Т-Comm - телекоммуникации и транспорт. – 2008. – № 6. – С. 21-23.
3. Поляков Д.Б. Методы пространственно-временного шумоподавления // Труды 14-й Всероссийской научно-технической конференции «Современное телевидение». – М.: ФГУП МКБ «Электрон», 2006. – С. 87-89.
4. Поляков Д.Б. Пространственно-временные искажения сигналов изображений со сжатым спектром // Научная конференция профессорско-преподавательского, научного и инженерно-технического состава: Тез. докл. Книга 1. – М., 2006. – С. 135-136.
5. Поляков Д.Б. Повышение точности оценки движения динамических объектов изображения // Материалы 5-й Международной научно-технической конференции Intermatic «Фундаментальные проблемы радиоэлектронного приборостроения». – М.: МИРЭА, 2006. – Ч.2. – С. 184-186.
6. Поляков Д.Б. Методы компенсации пространственно-временных искажений в системах со сжатием спектра сигналов изображений по стандарту MPEG-2 // Московская отраслевая научно-техническая конференция «Технологии информационного общества»: Тез. докл. – М.: Инсвязьиздат, 2007. – С. 120.
7. Поляков Д.Б. Специфика слежения за движением в стандарте MPEG-4 // Московская отраслевая научно-техническая конференция «Технологии информационного общества»: Тез. докл. – М.: Инсвязьиздат, 2007. – С. 121.
8. Поляков Д.Б. Блочные алгоритмы оценки движения // Труды Московского технического университета связи и информатики: – М.: «ИД Медиа Паблшер», 2008. – Т.1. – С. 463-466.
9. Поляков Д.Б. Алгоритмы поиска границ для полутоновых изображений // Труды 16-й Международной научно-технической конференции «Современное телевидение». – М.: ФГУП МКБ «Электрон», 2008. – С. 183-186.
10. Поляков Д.Б., Игнатов Ф.М. Адаптивное добавление шума в изображение для устранения блочных искажений // Труды Российского научно-технического сообщества радиотехники, электроники и связи им. А.С.

Попова / Серия: Научная сессия, посвященная Дню Радио / Выпуск: LXIII. – М.: Инсвязьиздат, 2008. – С. 359-361.

11. Поляков Д.Б. Исследование фильтров интерполяции в стандартах видеокодирования // Журнал научных публикаций аспирантов и докторантов. – 2008. – № 8. – С. 246-249.
12. Поляков Д.Б. Оценка качества фильтров интерполяции в стандартах видеокодирования // Материалы Международной научно-технической школы-конференции «Молодые ученые – науке, технологиям и профессиональному образованию». – М.: Энергоатомиздат, 2008. – С. 87-91.
13. Поляков Д.Б. Алгоритм селективного взвешенного предсказания в видеокодировании // Материалы Международной научно-технической школы-конференции «Молодые ученые – науке, технологиям и профессиональному образованию». – М.: Энергоатомиздат, 2008. – С. 146-149.