

УЧРЕЖДЕНИЕ ОБРАЗОВАНИЯ «БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ»

УДК 621.397.424:004.056.55

АЛЬ-ДЖУБУРИ
Тарик Мохаммед Сальман

**ЭФФЕКТИВНОЕ КОДИРОВАНИЕ ВИДЕОДАНЫХ В МОБИЛЬНЫХ
СИСТЕМАХ НА ОСНОВЕ ОБЪЕКТНОЙ КОМПЕНСАЦИИ
ДВИЖЕНИЯ КАМЕРЫ**

Автореферат диссертации на соискание ученой степени
кандидата технических наук
по специальности 05.12.13 – Системы, сети
и устройства телекоммуникаций

Минск, 2011

Работа выполнена в учреждении образования «Белорусский государственный университет информатики и радиоэлектроники»

Научный руководитель

Конопелько Валерий Константинович, доктор технических наук, профессор, кафедра сетей и устройств телекоммуникаций, учреждение образования «Белорусский государственный университет информатики и радиоэлектроники».

Тузиков Александр Васильевич, доктор физико-математических наук, профессор, Государственное научное учреждение «Объединенный институт проблем информатики Национальной академии наук Беларуси»;

Хижняк Александр Вячеславович, кандидат технических наук, доцент, кафедра автоматических систем управления войсками, учреждение образования «Военная академия Республики Беларусь».

Оппонирующая организация

Государственное научное учреждение «Физико-технический институт Национальной академии наук Беларуси».

Защита состоится 1 марта 2012 года в 14:00 на заседании совета по защите диссертаций Д 02.15.02 при учреждении образования «Белорусский государственный университет информатики и радиоэлектроники», по адресу 220013, г. Минск, ул. П. Бровки, 6, ауд. 232-1, e-mail: dissovet@bsuir.by, тел.: 293-89-89.

КРАТКОЕ ВВЕДЕНИЕ

Эффективное кодирование находит широкое применение в телекоммуникациях для сжатия информации, в том числе при передаче видеоданных в мобильных системах, благодаря чему имеется возможность повысить пропускную способность каналов связи. В настоящее время наибольшее распространение получили методы эффективного кодирования на основе блочной компенсации движения (MPEG-4, H.264), позволяющие сжимать видеоданные в десятки – тысячи раз за счет предсказания кадров. Методы блочной компенсации движения показывают хорошие результаты в случае неподвижной видеокамеры. Однако, при движении видеокамеры использование видеокодеков с блочной компенсацией движения не эффективно, поскольку из-за параллакса (различного смещения изображений объектов на соседних кадрах при движении видеокамеры в зависимости от расстояний до объектов) наблюдаются ошибки предсказания прогнозного кадра и снижение коэффициентов сжатия. В случае ограниченной полосы передачи видеоданных по радиоканалу это приводит к резкому ухудшению качества восстановления кадров. Необходимость получения высокого коэффициента сжатия в условиях параллакса из-за движения видеокамеры обуславливает актуальность темы диссертации. В работе исследуются методы и средства эффективного кодирования видеоданных в мобильных системах на основе объектной компенсации движения камеры, учитывающие параллакс при формировании прогнозного кадра.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Связь работы с крупными научными программами

Исследования проводились в рамках выполненной на кафедре сетей и устройств телекоммуникаций учреждения образования «Белорусский государственный университет информатики и радиоэлектроники» научно-исследовательской госбюджетной темы ГБ 06-2033 «Разработка методов обработки, передачи и распределения мультимедийной информации» (Этап «Разработать и исследовать высокоскоростные методы передачи мультимедийной информации по различным каналам связи»).

Цель и задачи исследования

Цель настоящей работы состоит в разработке методов и средств объектной компенсации движения камеры для эффективного кодирования видеоданных в мобильных системах.

Для достижения поставленной цели необходимо решить следующие задачи:

- разработать метод объектной компенсации движения видеокамеры на

основе объектной декомпозиции опорного кадра и информации о параметрах установки и движения видеокамеры;

- разработать метод монокулярной объектной декомпозиции изображений;
- разработать методы сегментации сложных структурно-текстурных изображений с большим числом деталей, обеспечивающие высокую стабильность результатов сегментации в условиях изменения яркости и контраста;

- разработать структуру и аппаратно-программные средства эффективного кодирования видеоданных с объектной компенсацией движения видеокамеры.

Объектами исследования настоящей работы являются методы и средства эффективного кодирования видеоданных в мобильных системах. Выбор данных объектов исследования обусловлен актуальностью проблемы повышения пропускной способности мобильных систем телекоммуникаций, эффективным подходом к решению которой является сокращение объема передаваемой информации за счет сжатия видеоданных. Предметом исследования являются методы и средства объектной компенсацией движения видеокамеры для эффективного кодирования видеоданных.

Положения, выносимые на защиту

1 Метод объектной компенсации движения видеокамеры на основе двумерной объектной декомпозиции видеоданных, обеспечивающий уменьшение ошибки предсказания до 4,5 раз и увеличение коэффициента сжатия разности предсказанного и прогнозируемого кадров до 2,5 раз в сравнении с методом блочной компенсации движения, используемым в видеокодеке MPEG-4.

2 Метод монокулярной объектной декомпозиции видеоданных с предсказанием положения линии горизонта, обеспечивающий уменьшение вычислительной сложности объектной декомпозиции в 5 раз по сравнению с методом на основе базовых линий.

3 Метод сегментации изображений на основе реверсивной кластеризации, позволяющий исключить ошибки сегментации сложных по структуре областей, повысить устойчивость результатов сегментации на 5 – 20% по сравнению с методом разделения и слияния областей.

4 Метод адаптивной сегментации полутоновых изображений в области многоуровневого дискретного вейвлет-преобразования на основе иерархического наращивания областей, позволяющий повысить устойчивость результатов сегментации до 8 раз по сравнению с методом наращивания областей.

Личный вклад соискателя

Содержание диссертации отображает личный вклад автора. Он заключается в научном обосновании методов объектной компенсации движения видеокамеры для эффективного кодирования видеоданных, постановке и проведении экспериментов по исследованию характеристик, оценке эффективности разработанных методов, обработке и анализе полученных результатов, формулировке выводов.

Определение целей и задач исследований, интерпретация и обобщение полученных результатов проводились совместно с научным руководителем д.т.н. В.К. Конопелько и к.т.н. В.Ю. Цветковым.

Апробация результатов диссертации

Основные положения и результаты диссертационной работы докладывались и обсуждались на следующих конференциях: Международной научно-технической конференции, посвященной 45-летию МРТИ–БГУИР (БГУИР, Минск, Беларусь, 2009 г.); First Computer Science Conference (University Of Technology, Baghdad, Iraq, 2010); Second Scientific conference for information technology «Applications and Horizons» (University Of Technology, Baghdad, Iraq, 2010); 5-ой Международной научной конференции по военно-техническим проблемам, проблемам обороны и безопасности, использованию технологий двойного применения (БЕЛИСА, Минск, Беларусь, 2011 г.); VIII, IX Белорусско-российской научно-технической конференции «Технические средства защиты информации» (БГУИР, Минск, Беларусь, 2010, 2011 г.); Международном научно-техническом семинаре «Телекоммуникации: Сети и технологии, алгебраическое кодирование и безопасность данных» (БГУИР, Минск, Беларусь, 2010, 2011 г.).

Опубликованность результатов диссертации

По результатам исследований, представленных в диссертации, опубликовано 14 печатных работ, в том числе: 5 статей в научных журналах общим объемом 2.2 авторских листа; 8 статей и тезисы в сборниках и материалах конференций. Получено положительное решение на выдачу патента.

Структура и объем диссертации

Работа состоит из введения, четырех глав, заключения, списка использованных источников и приложений. В первой главе рассмотрены особенности мобильных систем и представлен анализ методов эффективного кодирования, сегментации и объектной декомпозиции видеоданных. Во второй

главе приведены исследования влияния параллакса на эффективность компенсации движения, предложены методы объектной компенсации движения на основе двумерной объектной декомпозиции видеоданных и монокулярной объектной декомпозиции изображений. В третьей главе рассматриваются методы сегментации изображений на основе реверсивной кластеризации и иерархического наращивания областей. В четвертой главе рассматриваются вопросы организации, принципы работы и функционирование разработанного видеокodeка с объектной компенсацией движения, оценки его характеристик, приводятся рекомендации по выбору параметров основных узлов видеокodeка. Общий объем диссертационной работы составляет 127 страниц, из которых 74 страницы текста, 62 рисунка на 25 страницах, 4 таблицы на 1 странице, 6 приложений на 16 страницах, библиография из 120 источников на 11 страницах, включая 14 собственных публикаций автора и положительное решение на выдачу патента на 2 страницах.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении определены основные направления исследований, обоснована актуальность темы диссертации, показана необходимость разработки и исследования методов и средств эффективного кодирования видеоданных на основе блочной компенсации движения видеокамеры.

В первой главе проведен анализ особенностей построения мобильных систем, рассмотрены методы эффективного кодирования видеоданных на основе блочной компенсации движения и используемые в них методы сегментации и объектной декомпозиции.

Показано, что при видео кодировании в мобильных системах необходимо учитывать влияние скорости и направления движения камеры, глубины сцены, освещенности объектива на корреляцию соседних кадров видеоданных. Известные методы эффективного кодирования (рисунок 1), используемые в стационарных системах передачи, основаны на покадровом кодировании (MotionJPEG, MotionJPEG 2000), кадровой разности (H.263, Суперак), блочной компенсации движения (MPEG-2, MPEG-4, H.264) (рисунок 2) и не учитывают данные особенности.

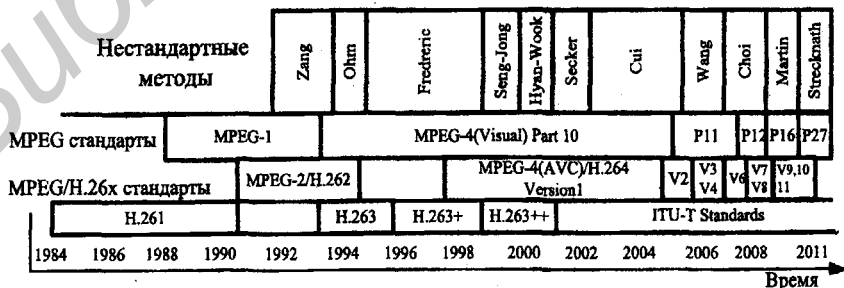


Рисунок 1 – Методы эффективного кодирования видеоданных

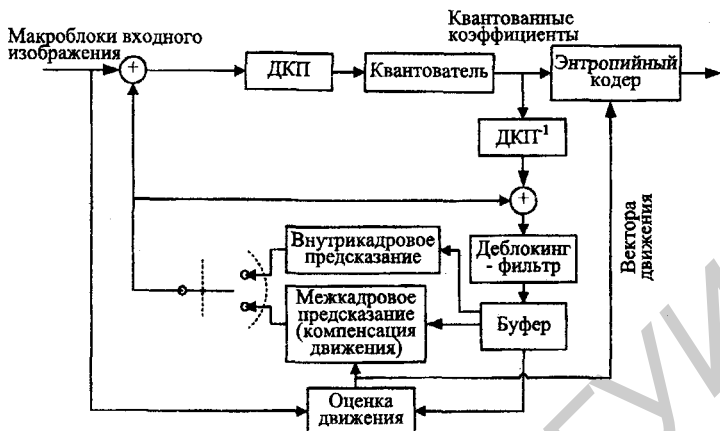


Рисунок 2 – Структурная схема видеокодера с блочной компенсацией движения

В результате ошибка предсказания текущего кадра по опорному оказывается большой, что не позволяет достичь высоких коэффициентов сжатия. Показана необходимость разработки методов эффективного кодирования видеоданных с объектной компенсацией движения камеры, позволяющих уменьшить ошибку предсказания и повысить коэффициент сжатия видеоданных в условиях параллакса для мобильных систем. Установлено, что выделение объектов на изображении основано на бинокулярной оптической модели в условиях известной стереобазы. Показано, что недостатком данного подхода является высокая вычислительная сложность. Отмечается, что вопросы использования монокулярной оптической модели для объектной декомпозиции и реализации объектной компенсации движения видеоканеры на ее основе детально не исследованы. И, кроме того, существующие методы объектной декомпозиции видеоданных, основанные на сегментации неподвижных изображений, не обеспечивают достаточной точности. Показана необходимость разработки методов сегментации, обеспечивающих высокую точность объектной декомпозиции, устойчивость результатов к изменению яркости и контраста изображения. Повысить устойчивость возможно за счет предварительной обработки на основе низкочастотной фильтрации. Основным недостатком подобных решений является высокая вычислительная сложность, снижение точности сегментации из-за потери мелких и искажения крупных деталей. Устранение данных недостатков возможно за счет использования для сегментации изображений совокупности их вейвлет-образов. Показана необходимость исследований по вейвлет-сегментации видеоданных.

Во второй главе представлены результаты исследования влияния параллакса на эффективность кодирования видеоданных, подтверждающие неэффективность методов видеокодирования на основе блочной компенсации движения в условиях

мобильных систем; предложены метод и алгоритм объектной компенсации движения видеокамеры для эффективного кодирования видеоданных [3-А, 6-А, 7-А, 8-А 12-А, 13-А,].

Произведена оценка изменений видеоданных, обусловленных параллаксом. Для этого использованы 45 полутоновых ландшафтных изображений с различных ракурсов (9 угловых ориентаций в горизонтальной плоскости с интервалом в 5° для 5 смещений камеры относительно оси вращения с интервалом 10 см). Проекционные искажения двух изображений оцениваются в области перекрытия с помощью среднеквадратической ошибки MSE, нормированной по яркости и площади области перекрытия изображений тестового набора и вычисляемой с помощью выражения

$$MSE = \sum_{y=y_{\min}}^{y_{\max}-1} \sum_{x=x_{\min}}^{x_{\max}-1} (\bar{m}_F(y, x, \varphi_i, s_i) - \bar{m}_F(y, x, \varphi_j, s_j))^2 / ((x_{\max} - x_{\min})(y_{\max} - y_{\min})), \quad (1)$$

где y_{\max} , y_{\min} , x_{\max} , x_{\min} – координаты области перекрытия изображений; i, j – номера изображений; φ_i, φ_j – угловые ориентации камеры в горизонтальной плоскости для изображений i и j ; s_i, s_j – смещения камеры относительно оси вращения в горизонтальной плоскости для изображений i и j ; $\bar{m}_F(y, x, \varphi_i, s_i)$, $\bar{m}_F(y, x, \varphi_j, s_j)$ – нормированные по яркости пиксели области перекрытия изображений i и j .

На рисунке 3 представлены зависимости ненормированной и нормированной по яркости среднеквадратической ошибки MSE областей перекрытия тестового изображения, полученного на нулевой угловой ориентации видеокамеры в горизонтальной плоскости при смещении видеокамеры на 20 см ($\varphi_i = 0^\circ$, $s_i = 20$ см), и каждого из остальных 44 тестовых изображений ($\varphi_j \in [0^\circ, 40^\circ]$, $s_j \in [0, 40]$, $(\varphi_i \neq \varphi_j) \wedge (s_i \neq s_j)$) при смещении оптической оси камеры относительно оси вращения в горизонтальной плоскости на 10 см (кривая s10), 20 см (кривая s20), 30 см (кривая s30) и 40 см (кривая s40). Значения на оси абсцисс соответствуют номерам разностных изображений.

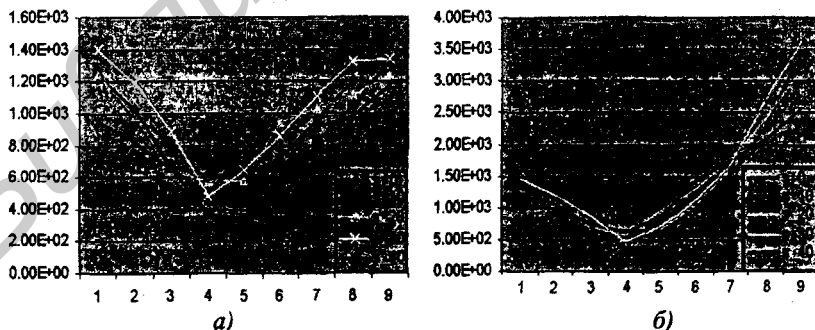


Рисунок 3 – Зависимости MSE от угла поворота и смещения видеокамеры: а – ненормированной; б – нормированной по яркости

Показано, что проекционные искажения, обусловленные параллаксом, возрастают примерно в 3 раза при смещении видеокамеры на 20° . Полученные результаты свидетельствуют о неэффективности устранения временной избыточности видеоданных при движении видеокамеры на основе кадровой разности и блочной компенсации движения.

Предложен метод объектной компенсации движения видеокамеры, основанный на двухмерной объектной декомпозиции и объектном предсказании прогнозного кадра видеоданных с использованием опорного кадра и параметров установки и перемещения камеры. Метод позволяет устранять временную избыточность при кодировании видеоданных в условиях движения камеры. Пиксельная разность кадра-прогноза и реального кадра состоит в основном из нулевых или близких к нулю значений и может быть эффективно закодирована. Реализация метода предполагает выполнение следующих операций.

1. Двухмерная объектная декомпозиция опорного кадра $M_F(t)$ видеоданных ($t \in [0, T-1]$) размером $Y \times X$ пикселей. Формируется матрица M_D объектной декомпозиции, содержащая информацию о количестве N_O и расположении объектов опорного кадра $M_F(t)$, которые представляются совокупностями одинаковых по значениям пикселей.

2. Определение расстояний от камеры до объектов опорного кадра $M_F(t)$ видеоданных. Формируется матрица D_Z расстояний от камеры до элементов трехмерной сцены, проецируемых на соответствующие пиксели опорного кадра.

3. Объектное предсказание изменений на новом кадре $M_F(t+i)$ видеоданных ($i > 0$). Формируется прогнозный кадр $\hat{M}_F(t+i) = \|\hat{m}_F(y, x, t+i)\|_{(y=0, Y-1, x=0, X-1)}$ для нового положения камеры в момент $t+i$. При перемещении камеры объекты опорного кадра $M_F(t)$ сдвигаются и масштабируются на прогножном кадре $\hat{M}_F(t+i)$ в зависимости от их удаленности от камеры. Смещения и изменения размеров объектов на прогножном кадре вычисляются с помощью модели $M_C(t)$ монокулярного дальномера.

4. Вычисление ошибки предсказания. На основе прогнозного $\hat{M}_F(t+i)$ и реального $M_F(t+i)$ кадров формируется разностный кадр $\tilde{M}_F(t+i)$.

5. Сжатие разностного кадра $\tilde{M}_F(t+i)$.

Предложен алгоритм объектного предсказания изменения видеоданных при движении камеры по опорному кадру, соответствующему ее исходному положению, и параметрам ее перемещения (рисунок 4). Алгоритм основан на модели движения камеры, учитывающей изменение расстояний до объектов сцены при произвольном перемещении камеры, и модели монокулярного дальномера, позволяющей определять расстояние до объекта сцены по смещению проекции точки его местоположения на изображении и известным параметрам установки камеры. В результате выполнения алгоритма формируется прогнозный кадр, соответствующий новому положению видеокамеры при ее произвольном перемещении.

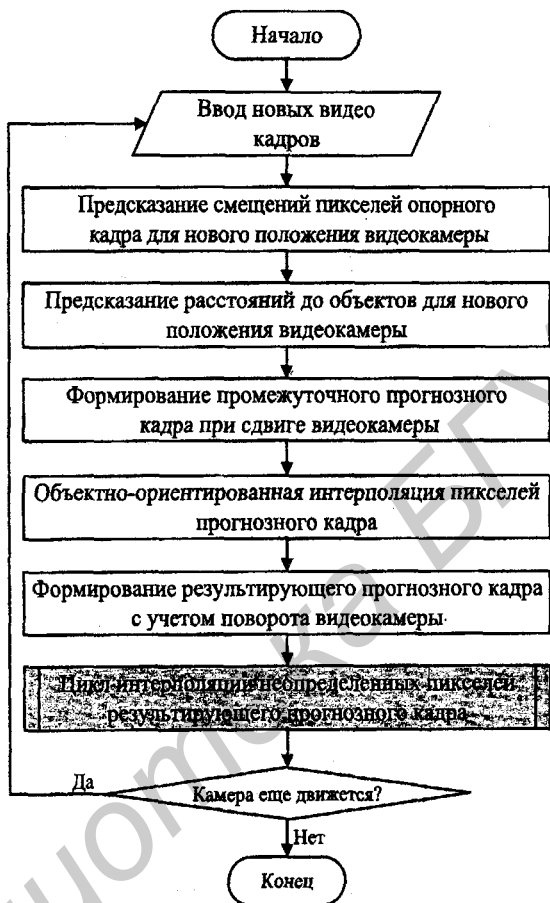


Рисунок 4 – Алгоритм объектного предсказания изменения видеоданных при движении видеокамеры

В таблице 1 представлены значения среднеквадратичной ошибки (MSE) предсказания изменений на изображении и коэффициенты (CR) сжатия без потерь (архиватор WinRar) разности прогнозного и реального нового кадров видеоданных при горизонтальном (HM) и вертикальном (VM) движении камеры на 1, 2, 3 и 4 см для трех методов устранения временной избыточности: Cinepak, MPEG-4 и объектной компенсации движения камеры на основе двухмерной декомпозиции видеоданных (ОКД). Данные получены для полутоновых опорных кадров с несколькими объектами общей площадью порядка 1/8 площади изображения.

Таблица 1– Характеристики методов предсказания изменений видеоданных

Значения сдвига видеокамеры относительно исходного положения	Методы устранения временной избыточности видеоданных											
	Синерак				MPEG-4				ОКД			
	MSE		CR		MSE		CR		MSE		CR	
	HM	VM	HM	VM	HM	VM	HM	VM	HM	VM	HM	VM
1 см	2362	3231	2.0	2.0	1294	1541	5.1	5.9	215	513	9.0	13.4
2 см	3604	1770	2.0	2.0	1868	1611	4.1	3.8	615	537	10.4	9.1
3 см	3713	3904	2.5	1.6	2202	1842	3.2	4.8	623	603	7.7	8.4
4 см	3801	2966	2.2	2.1	2802	2202	4.8	4.8	624	731	9.9	6.7

Показано, что метод объектной компенсации движения камеры обеспечивает уменьшение ошибки предсказания примерно в 6 раз по сравнению с методом кадровой разности, используемом в Синерак, и примерно в 4,5 раза по сравнению с методом блочной компенсации движения, используемом в MPEG-4. Метод обеспечивает выигрыш в коэффициенте сжатия без потерь разности предсказанного и реального прогнозируемого кадров видеоданных до 5 раз в сравнении с методом Синерак и до 2,5 раз в сравнении с методом MPEG-4.

Предложен метод монокулярной объектной декомпозиции изображения, основанный на предсказании положения линии горизонта на изображении, позволяющий отделять объекты от фона. В данном методе сегменты классифицируются как изображения объектов, если для них выполняется условие

$$(h_y(n_s) > Y_G - \Delta_G) \vee (l_y(n_s) < Y_G + \Delta_G), \quad (2)$$

где Δ_G – задаваемое допустимое отклонение положения проекции линии горизонта; $h_y(n_s)$, $l_y(n_s)$ – y-координаты верхних и нижних пикселей сегмента n_s ; Y_G – смещение линии горизонта относительно нижней границы выровненного сегментированного опорного кадра. Сегменты, не удовлетворяющие данному выражению, относятся к фону. Вычислительная сложность предложенного метода монокулярной объектной декомпозиции на основе предсказания положения линии горизонта примерно в 5 раз ниже по сравнению с методом на основе базовых линий, за счет обработки 1 кадра вместо 5.

В третьей главе предложены методы и алгоритмы сегментации изображений на основе реверсивной кластеризации и иерархического наращивания областей, необходимые для объектной декомпозиции опорных кадров видеоданных и последующей объектной компенсации движения видеокамеры [1-А, 4-А, 5-А, 9-А, 10-А, 11-А].

Сущность метода RCPS (Reversible Cluster Progressive Segmentation) сегментации на основе реверсивной кластеризации состоит в выделении однородных по яркости областей пикселей с помощью компактного описания структуры изображения с помощью карт кластеризации однородностей, аппроксимации и сегментации. Карты являются многоуровневыми, содержат на каждом уровне кратномасштабное представление изображения и

обрабатываются реверсивно – сначала формируются карты кластеризации и аппроксимации (прямой проход), а затем с их помощью формируется карта сегментации (обратный проход). При этом обеспечивается устранение ошибок сегментации сложных по структуре областей за счет верификации компактных описаний соседних сегментов на каждом уровне кластеризации. Алгоритм, реализующий данный метод, включает выделение областей на основе итеративной кластеризации (прямой проход) и прогрессивную сегментацию на основе разделения, слияния и поглощения областей (обратный проход).

Кластеризация основана на формировании матрицы $A(l)$ аппроксимации l -го уровня, элементы которой вычисляются с помощью выражения

$$a^{(l)}(y, x) = \begin{cases} 0 & \text{при } c^{(l)}(y, x) = 1, \\ S_A & \text{при } c^{(l)}(y, x) = 0 \end{cases} \quad (3)$$

при $y = \overline{0, Y/2^l - 1}$, $x = \overline{0, X/2^l - 1}$, где $S_A = \frac{1}{4} \sum_{j=0}^1 \sum_{i=0}^1 a^{(l-1)}(2y+j, 2x+i)$ – сумма элементов кластера с координатами $(2y, 2x)$ в матрице $A(l-1)$ аппроксимации $l-1$ уровня; $c^{(l)}(y, x)$ – элементы матрицы кластеризации. Для определения элементов матрицы $C(l)$ кластеризации l -го уровня используется выражение

$$c^{(l)}(y, x) = \begin{cases} 0 & \text{при } (S_C = 0) \wedge (D_A \leq T_S), \\ 1 & \text{при } (S_C > 0) \vee (D_A > T_S) \end{cases} \quad (4)$$

при $y = \overline{0, Y/2^l - 1}$, $x = \overline{0, X/2^l - 1}$, где $S_C = \sum_{j=0}^1 \sum_{i=0}^1 c^{(l-1)}(2y+j, 2x+i)$ – сумма элементов кластера с координатами $(2y, 2x)$ в матрице $C(l-1)$ кластеризации нижнего $(l-1)$ -го уровня;

$$D_A = \frac{\left(\max_{j=0}^1 \left(\max_{i=0}^1 (a^{(l-1)}(2y+j, 2x+i)) \right) - \min_{j=0}^1 \left(\min_{i=0}^1 (a^{(l-1)}(2y+j, 2x+i)) \right) \right)}{\max_{j=0}^1 \left(\max_{i=0}^1 (a^{(l-1)}(2y+j, 2x+i)) \right)} - \text{взвешенная}$$

разность между максимальным и минимальным значениями элементов кластера с координатами $(2y, 2x)$ в матрице $A(l-1)$ аппроксимации нижнего $(l-1)$ -го уровня.

При прогрессивной сегментации формируются значения элементов матрицы $S(l-1)$ сегментации $(l-1)$ -го уровня с помощью выражения (масштабирование областей)

$$(c^{(l)}(y, x) = 0) \rightarrow \forall j (j = \overline{0, 1}) \forall i (i = \overline{0, 1}) (s^{(l-1)}(2y+j, 2x+i) = s^{(l)}(y, x)) \quad (5)$$

при $y = \overline{0, Y/2^l - 1}$, $x = \overline{0, X/2^l - 1}$. Прогрессивная сегментация включает формирование новых областей (разделение областей), наращивание областей путем их добавления к существующим однородным областям и обработку конфликтующих номеров сегментов.

Показано, что метод RCPS обеспечивает точную локализацию областей любой сложности. Наиболее близкими к нему по принципам организации и функционирования являются метод разделения и слияния областей на основе квадрата-дерева CQTS и его модификация MQTS. Для методов RCPS и MQTS на рисунке 5 представлены зависимости числа сегментов, выделяемых на тестовых изображениях, от изменения яркости и контраста, из которых следует, что метод RCPS более устойчив к изменению яркости по сравнению с методом MQTS (для большинства изображений устойчивость метода RCPS выше на 5 – 20 %).

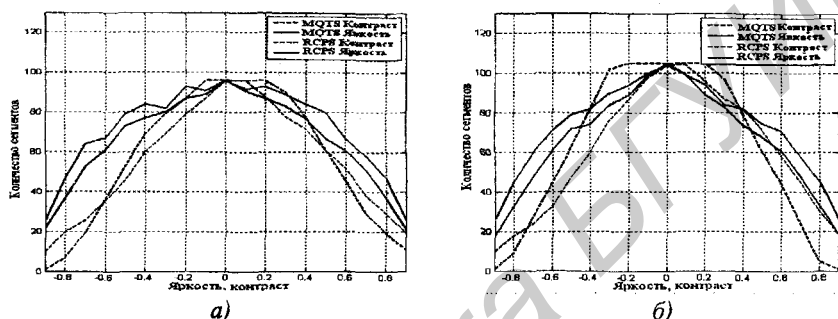


Рисунок 5 – Зависимости числа сегментов от яркости и контраста изображений: *a* – Barbara; *б* – Lena

Предложен метод HRG (Hierarchic Region Growing) адаптивной сегментации полутоновых изображений в области многоуровневого дискретного рационального вейвлет-преобразования, обеспечивающий устойчивость результатов выделения элементов на сложных структурно-текстурных изображениях с большим числом деталей, основанный на иерархическом наращивании областей, сегментированных на верхнем уровне вейвлет-декомпозиции. В данном методе формируется матрица $M_R(l) = \|m_R^{(l)}(y, x)\|_{\left(\substack{y=0, Y/2^l-1, \\ x=0, X/2^l-1}\right)}$ увеличенной (масштабированной) проекции верхнего уровня сегментации на нижний, элементы которой вычисляются с помощью выражения

$$m_R^{(l)}(y, x) = m_R^{(l+1)}\left(\left\lfloor \frac{y}{2} \right\rfloor, \left\lfloor \frac{x}{2} \right\rfloor\right) \quad (6)$$

при $y = \overline{0, Y/2^l - 1}$ и $x = \overline{0, X/2^l - 1}$, где $\lfloor \cdot \rfloor$ – операция округления с недостатком;

$m_R^{(l+1)}\left(\left\lfloor \frac{y}{2} \right\rfloor, \left\lfloor \frac{x}{2} \right\rfloor\right)$ – элементы матрицы $M_R(l+1)$ наращивания областей $(l+1)$ -го уровня, которые формируются на основе элементов матрицы $M_R(l)$ сегментации L -го уровня с помощью выражения

$$m_R^{(L)}(y, x) = m_B^{(L)}(y, x) \quad (7)$$

при $y = 0, Y/2^L - 1$ и $x = 0, X/2^L - 1$.

Для оценки эффективности метода HRG использованы полутоновые изображения различных классов (низкочастотные и высокочастотные) размером 128×128 пикселей. Сравнение метода HRG с известным методом наращивания областей RG произведено по площадям (в пикселях) сегментов, выделенных на тестовых изображениях в условиях изменения их яркости Y и контраста C (рисунок 6). Оцениваются площади $A_R^{(HRG)}$ и $A_R^{(RG)}$ одного из сегментов, выделенных на одном изображении с помощью адаптивной вейвлет-сегментации и наращивания областей соответственно.

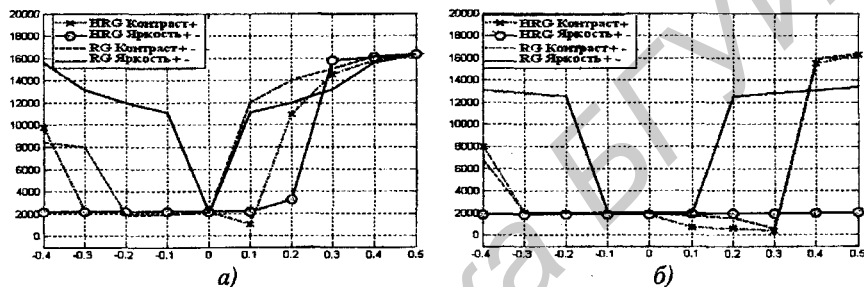


Рисунок 6 – Зависимости площади (ось ординат) от изменения яркости и контраста изображений: а – Lena; б – car

Для оценки устойчивости $S_Y(A_r)$ результатов сегментации к изменению яркости используется выражение

$$S_Y(A_r) = \frac{\Delta A_r(\Delta Y)}{\Delta Y}, \quad (8)$$

где $\Delta Y = Y_{MAX} - Y_{MIN}$ – интервал изменения яркости, на котором производится анализ устойчивости результатов сегментации; Y_{MAX} , Y_{MIN} – максимальные и минимальные значения яркости, определяющие интервалы анализа; $A_r(\Delta Y) = A_{rMAX}(\Delta Y) - A_{rMIN}(\Delta Y)$ – диапазон изменения площади A_r контролируемого сегмента на интервале ΔY .

Показано, что разработанный метод адаптивной вейвлет-сегментации на основе иерархического наращивания областей обеспечивает повышение устойчивости результатов сегментации к изменению яркости до 8 раз по отношению к методу сегментации на основе наращивания областей. Чувствительность обоих методов к изменению контраста приблизительно одинакова. Устойчивость результатов сегментации в методе HRG достигается за счет многоуровневой аппроксимации на основе вейвлет-преобразования, сглаживающей мелкие детали и текстурные особенности сегментов.

В четвертой главе рассматриваются структура, принципы работы и функционирование, оценка характеристик и рекомендации по выбору параметров видеокодека с объектной компенсацией движения видеокамеры [1-А, 2-А, 3-А, 4-А, 5-А, 14-А]. На рисунке 7 представлена структурная схема видеокодека с объектной компенсацией движения камеры. Видеокодек обеспечивает сжатие видеоданных, поступающих от движущейся камеры, учитывает параллакс, и позволяет по сравнению с видеокодеками на основе блочной компенсации движения и кадровой разности повысить коэффициент сжатия и качество восстановления видеоданных в условиях движения камеры.

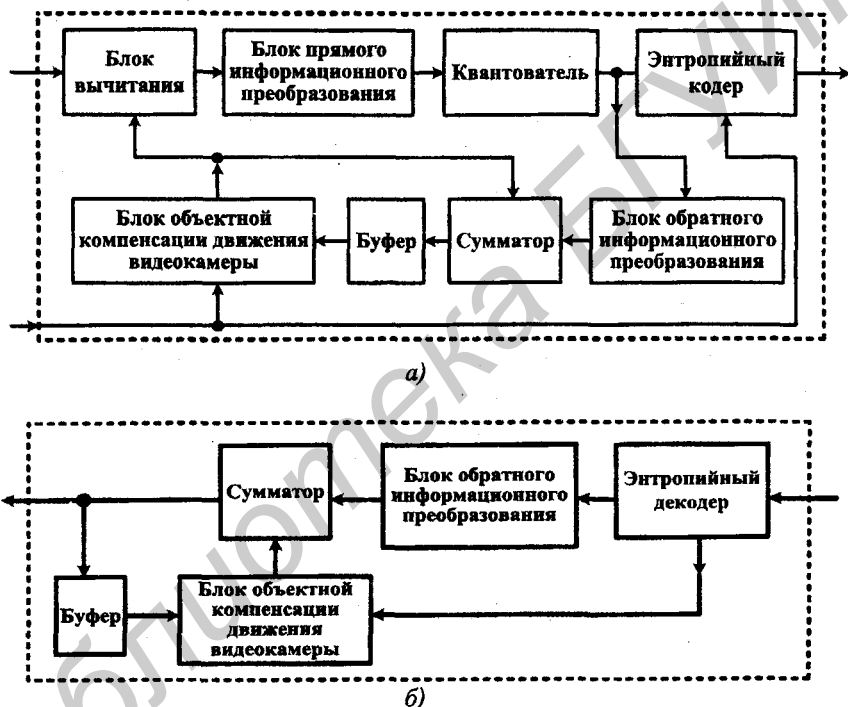


Рисунок 7 – Видеокодек с объектной компенсацией движения видеокамеры:
 а – видеокодер; б – видеодекoder

Для оценки представления информации о движении использован коэффициент K_C компактности, имеющий размерность бит/пиксель, показывающий отношение объема координатной информации V_M о движении к размеру кадра и определяемый выражением

$$K_C = V_M / (Y X), \quad (9)$$

где $V_M = Y \times X \times BD$; BD – битовая глубина.

Показано, что при использовании 8 бит на кодирование каждого из 7 параметров установки и движения видеокамеры ($V_M = 8 \times 7 = 56$ бит) и размера кадра 480×640 пикселей предложенный видеокодек обеспечивает значение коэффициента $K_C = 1,8 \cdot 10^{-4}$ бит/пиксель. Для сравнения, при использовании 16 бит на кодирование вектора движения каждого блока, размера кадра 480×640 пикселей и размера блока $8 \times 8 = 64$ пикселей ($V_M = 16 \times 480 \times 640 / 64 = 76800$ бит) видеокодек MPEG-4 обеспечивает значение коэффициента $K_C = 0,25$ бит/пиксель (примерно в 1400 раз хуже по сравнению с предложенным видеокодеком).

На рисунке 8 представлены зависимости коэффициентов сжатия для видеокодека MPEG-4 (CR_B) и видеокодека с объектной компенсацией движения (CR_O), определяемые как $V_F / (V_F / CR_F + V_M)$, от коэффициента CR_F сжатия кадра без учета информации о векторах движения камеры. Показано, что видеокодек с объектной компенсацией движения превосходит в единицы раз видеокодек MPEG-4. Выигрыш тем больше, чем больше размер кадра и коэффициент CR_F .

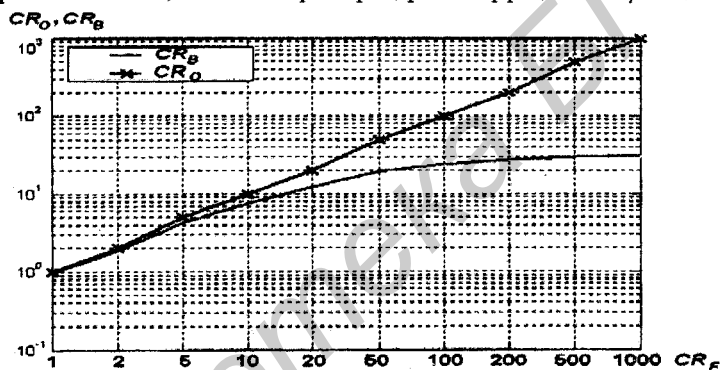


Рисунок 8 – Графики зависимости CR_B, CR_O от CR_F

Точность предсказания оценивается с использованием пикового отношения сигнал-шум $PSNR$, вычисляемого для области предсказания с помощью выражения

$$PSNR = 10 \log_{10} \left((2^{BD} - 1)^2 / MSE \right), \quad (10)$$

где $MSE = \frac{1}{YX} \sum_{y=0}^{Y_0-1} \sum_{x=0}^{X_0-1} (f(y, x, t) - \hat{m}_F(y, x, t))^2$ – среднеквадратическая ошибка; Y_0, X_0 – размеры области перекрытия.

В таблице 2 приведены значения пикового отношения $PSNR$ сигнал-шум предсказания, полученные при сжатии тестовой последовательности из 12 кадров с помощью разработанного видеокодека и видеокодека MPEG-4 при размере блока 8×8 пикселей. Показано, что предложенный видеокодек обеспечивает выигрыш в пиковом отношении $PSNR$ сигнал-шум предсказания по сравнению с видеокодеком MPEG-4.

Таблица 2 – Значения пикового отношения *PSNR* сигнал-шум предсказания, дБ

Номера кадров	1-2	3-4	5-6	7-8	9-10	11-12
Предложенный видеокодек	69,70	69,14	68,40	67,15	67,36	68,55
Видеокодек MPEG-4	68,17	67,47	67,52	65,42	67,02	68,20

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

1 Предложен метод объектной компенсации движения видеокамеры на основе двухмерной объектной декомпозиции видеоданных, позволяющий осуществить объектную декомпозицию опорного кадра видеоданных, определить расстояния до объектов опорного кадра, предсказать изменения на прогнозируемом кадре, вычислить и сжать ошибку предсказания. Установлено, что предложенный метод обеспечивает уменьшение ошибки предсказания до 4,5 раз и увеличение коэффициента сжатия разности предсказанного и прогнозируемого кадров до 2,5 раз по сравнению с методом блочной компенсации движения, используемым в видеокодеке MPEG-4 [2-А, 3-А, 7-А, 8-А].

2 Разработан метод монокулярной объектной декомпозиции изображения на основе предсказания положения линии горизонта, позволяющий отнести сегменты изображения к двум классам (объектам или фону) в зависимости от их размещения относительно предполагаемой линии горизонта. Показано, что данный метод обеспечивает уменьшение вычислительной сложности объектной декомпозиции в 5 раз по сравнению с методом на основе базовых линий [3-А, 7-А, 13-А].

3 Предложен метод сегментации изображений на основе реверсивной кластеризации, позволяющий исключить ошибки сегментации и последующей объектной декомпозиции сложных по структуре областей за счет использования процедуры верификации компактных описаний соседних сегментов на каждом уровне кластеризации. Установлено, что предложенный метод обеспечивает повышение устойчивости результатов сегментации на 5 – 20 % по сравнению с методом разделения и слияния областей [5-А, 9-А, 11-А].

4 Разработан метод адаптивной вейвлет-сегментации сложных структурно-текстурных изображений с большим числом деталей на основе иерархического наращивания областей, сегментированных на верхнем уровне вейвлет-декомпозиции, обеспечивающий устойчивость результатов сегментации и последующей объектной декомпозиции к изменению яркости

и контраста сегментируемого изображения [1-А, 4-А, 5-А, 10-А]. Показано, что данный метод в 8 раз более устойчив к изменению яркости по сравнению с методом сегментации на основе наращивания областей.

Рекомендации по практическому использованию результатов

1 Разработанные методы объектно-ориентированной компенсации движения видеокамеры, монокулярной объектной декомпозиции изображения и сегментации использованы при создании структуры видеокодека, алгоритмов и программных средств объектной компенсацией движения видеокамеры для эффективного кодирования видеоданных в мобильных системах. Показано, что при размере кадра 480×640 пикселей разработанные программные средства обеспечивают увеличение коэффициента сжатия примерно в 4 раза и пикового отношения сигнал-шум предсказания на величину до 1,5 дБ по сравнению с видеокодеком MPEG-4 за счет использования одного вектора движения для всего кадра вместо множества векторов движения для каждого блока и объектной компенсации движения вместо блочной. Установлено, что скорость предсказания в разработанном видеокодеке не ниже скорости предсказания видеокодека MPEG-4 [1-А, 2-А, 3-А, 4-А, 5-А, 6-А, 12-А, 14-А].

2 Показано, что при необходимости ускорения процесса предсказания изменений опорного кадра в разработанном видеокодеке для вейвлет-сегментации на всех уровнях вейвлет-декомпозиции могут использоваться рациональные вейвлеты Хаара вместо рациональных биортогональных вейвлетов 9.7, что позволяет увеличить скорость сегментации примерно в 4 раза [1-А, 4-А, 5-А].

3 Установлено, что на точность монокулярной объектной декомпозиции изображения существенное влияние оказывают пороговые значения, используемые для предсказания положения линии горизонта. Показано, что для повышения точности монокулярной объектной декомпозиции необходимо уменьшать высоту установки видеокамеры. При уменьшении высоты установки видеокамеры в 2 раза допустимое отклонение положения линии горизонта уменьшается примерно во столько же раз [3-А, 7-А, 8-А, 13-А].

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ

Статьи

1–А. Борискевич, А.А. Вейвлет–сжатие полутоновых изображений на основе одномерных комбинированных древесных структур / А.А. Борискевич., В.Ю. Цветков, Т.М. Аль-Джубури, О.Дж. Аль-Фурайджи // Доклады БГУИР, – 2008. №7(37).– С.59–66.

2–А. Аль-Джубури, Т.М. Объектное предсказание изменения видеоданных при горизонтальном перемещении камеры / Т.М. Аль-Джубури, В.Ю. Цветков, В.К. Конопелько // Доклады БГУИР. – 2009. – №5(43). – С. 4–12.

3–А. Аль-Джубури, Т.М. Метод объектно-ориентированной компенсации движения камеры на основе двумерной декомпозиции видеоданных / Т.М. Аль-Джубури, В.К. Конопелько, В.Ю. Цветков // Инженерный вестник. 2010. – Т. 2(30). – № 30. С. 20–26.

4–А. Boriskevich, A.A. Combined Hierarchical Wavelet–Coefficient Structures For Grayscale Image Compression / A.A. Boriskevich, V. Yu. Tsviatkou, T.M. Al–Juboori // Engineering and technology journal, Scientific journal published by the university of technology, Baghdad–Iraq. – 2010. – Vol. 28, № 5. – P. 1024–1037.

5–А. Аль-Джубури, Т.М. Адаптивная вейвлет-сегментация изображений на основе иерархического выращивания областей / Т.М. Аль-Джубури, В.К. Конопелько, В.Ю. Цветков // Доклады БГУИР. – 2011. – №2(56). – С. 102–108.

Материалы конференций

6–А. Цветков, В.Ю. Метод формирования синтетического видео на основе движения камеры / В.Ю. Цветков, Т.М. Аль-Джубури // Международная научно-техническая конференция, посвященная 45-летию МРТИ-БГУИР: тез. докл., Минск, 19 марта 2009г. – БГУИР, Минск: 2009. – С. 41–42.

7–А. Al-Juboori, T.M. Data modification based on monocular optar model / T.M. Al-Juboori, V.K. Kanapelka, V.Yu. Tsviatkau // First Computer science conference proceedings computer sciences are the base of society information. 10–12 February 2010 Baghdad / University Of Technology, Iraq, Baghdad. – 2010. – P 216–228.

8–А. Al-Juboori, T.M. Object oriented compensation of camera motion in 3D space based on distance range calculations / T.M. Al-Juboori, V.K. Kanapelka, V.Yu. Tsviatkau // 2nd Scientific conference for information technology: Applications and Horizons, UOT, Iraq– Baghdad, Part 4, 20–22 April 2010. – P. 223–226.

9—А. Факир, М.М. Многоуровневая защита сегментированных изображений на основе неравномерного шифрования и перестановок кластерных блоков / М.М. Факир, Т.М. Аль-Джубури // Технические средства защиты информации: тез. докл. VIII Белорусско-российской науч.-техн. конф., Браслав, Беларусь, 24–28 мая 2010г. / Белорус. гос. ун-т информатики и радиоэлектроники; редкол.: Л.М. Лыньков [и др.]. – Минск: БГУИР, 2010. – С. 38.

10—А Аль-Джубури, Т.М. Прогрессивная сегментация аппроксимированных вейвлет-образов на основе разделения-слияния и выращивания областей / Т.М. Аль-Джубури, В.Ю. Цветков // Телекоммуникации: сети и технологии, алгебраическое кодирование и безопасность данных: тез. докл. Междунар. науч.-технич. семинара, Браслав, Беларусь, 22–24 сент. 2010г. / Белорус. гос. ун-т информатики и радиоэлектроники; редкол.: В.К. Конопелько [и др.]. – Минск: БГУИР, 2010. – С. 48–50.

11—А. Аль-Джубури, Т.М. Сегментация изображений на основе реверсивной кластеризации / Т.М. Аль-Джубури, В.К. Конопелько, В.Ю. Цветков // 5-я Международная научная конференция по военно-техническим проблемам, проблемам обороны и безопасности, использованию технологий двойного применения: тез. докл., Минск, 25–26 мая 2011г. – ГУ «БелИСА», 2011. – С. 87–89.

12—А. Аль-Джубури, Т.М. Видеокодирование с объектной компенсацией движения и неравномерной криптографической защитой / Т.М. Аль-Джубури, В.К. Конопелько // Технические средства защиты информации: тез. докл. IX Белорусско-российской науч.-техн. конф., Браслав, Беларусь, 28–29 июня 2010г. / Белорус. гос. ун-т информатики и радиоэлектроники; редкол.: Л.М. Лыньков [и др.]. – Минск: БГУИР, 2011. – С. 52.

13—А. Аль-Джубури, Т.М. Монокулярная объектная декомпозиция изображения на основе предсказания положения линии горизонта / Т.М. Аль-Джубури, В.Ю. Цветков // Телекоммуникации: сети и технологии, алгебраическое кодирование и безопасность данных: материалы докл. междунар. науч.-техн. семинара, Минск, январь-декабрь 2011г. / Белорус. гос. ун-т информатики и радиоэлектроники; редкол.: В.Ф. Голенков [и др.]. – Минск, 2011. – С. 62–66.

Патенты

14—А. Положительное решение на выдачу патента «Видеокодек с объектной компенсацией движения видекамеры для сжатия видеоданных» от 29.12.2011 / В.К. Конопелько, В.Ю. Цветков, Т.М.С. Аль-Джубури, О.Дж.М Аль-Фурайджи; заявитель: Белорус. гос. ун-т информатики и радиоэлектроники – Заявка № и 20110853 от 31.10.2011.



РЭЗЮМЭ

Аль-Джубуры Тарык Махаммед Сальман

ЭФЕКТЫЎНАЕ КАДАВАННЕ ВІДЭАДАДЗЕННЫХ У МАБІЛЬНЫХ СІСТЭМАХ НА АСНОВЕ АБ'ЕКТНАГА КАМПЕНСАВАННЯ РУХУ КАМЕРЫ

Ключавыя словы: эфектыўнае кадаванне відэададзеных, аб'ектнае кампенсаванне руху відэакамеры, спіск відэададзеных, манакулярная аб'ектная декампазіцыя, прадказанне змены відэададзеных.

Мэтай працы з'яўляецца распрацоўка метадаў і сродкаў аб'ектнага кампенсавання руху камеры для эфектыўнага кадавання відэададзеных у мабільных сістэмах маніторынгу.

Даследаван ўплыў паралаксу на эфектыўнасць кадавання відэададзеных, фармаваных калі рухаецца відэакамерай. Усталявана, што выкарыстанне метадаў кадавання відэададзеных на аснове кадравай рознасці і блокавага кампенсавання руху ва ўмовах паралаксу неэфектыўна з-за значнай памылкі прадказання кадраў. Для падвышэння каэфіцыента спіску відэададзеных, фармаваных пры руху камеры, распрацаваны метады аб'ектнага кампенсавання руху камеры на аснове двухмернай аб'ектнай декампазіцыі і алгарытм яго рэалізацыі на аснове манакулярнай аптычнай мадэлі. Іста метады складаецца ў сегментацыі і аб'ектнай декампазіцыі апорнага кадра відэададзеных і прадказанні чарговага кадра з выкарыстаннем інфармацыі аб усталёўцы і перасоўванні відэакамеры. У адрозненне ад вядомых метадаў эфектыўнага кадавання відэададзеных, выкарыстоўвалых блокавую сегментацыю, аддзяленне рухомах аб'ектаў ад фону і вектары руху для кожнага блока, распрацаваны метады выкарыстае аб'ектна-арыентаваную сегментацыю, якая ўлічвае структуру малюнка, поўную аб'ектную декампазіцыю, што адлучае нерухомыя аб'екты сябар ад сябра, і адзін вектар руху для прадказання. За рахунак гэтага забяспечваецца памяншэнне памылкі прадказання да 4,5 раз і павелічэнне каэфіцыента спіску рознасці прадказанага і прагназуемага кадраў да 2,5 раз у параўнанні з метадам блокавага кампенсавання руху, выкарыстоўваным у відэакодэку MPEG-4.

На аснове прапанаваных метадаў і алгарытмаў эфектыўнага кадавання відэададзеных распрацаваны праграмныя модулі відэакодэка з аб'ектным кампенсаваннем руху камеры для спіску відэададзеных у мабільнай сістэме маніторынгу і павелічэнні за рахунак гэтага яе прапусканой здольнасці. Вынікі дысертацыйнай працы выкарыстаныя ў КБ «Дысплей» і ўкараненыя ў вучэбны працэс у Беларускай дзяржаўнай універсітэце інфарматыкі і радыёэлектронікі.

РЕЗЮМЕ

Аль-Джубури Тарик Мохаммед Сальман

ЭФФЕКТИВНОЕ КОДИРОВАНИЕ ВИДЕОДАНЫХ В МОБИЛЬНЫХ СИСТЕМАХ НА ОСНОВЕ ОБЪЕКТНОЙ КОМПЕНСАЦИИ ДВИЖЕНИЯ КАМЕРЫ

Ключевые слова: эффективное кодирование видеоданных, объектная компенсация движения видеокамеры, сжатие видеоданных, монокулярная объектная декомпозиция, предсказание изменения видеоданных.

Целью работы является разработка методов и средств объектной компенсации движения камеры для эффективного кодирования видеоданных в мобильных системах.

Исследовано влияние параллакса на эффективность кодирования видеоданных, формируемых движущейся видеокамерой. Установлено, что использование методов кодирования видеоданных на основе кадровой разности и блочной компенсации движения в условиях параллакса неэффективно из-за значительной ошибки предсказания кадров. Для повышения коэффициента сжатия видеоданных, формируемых при движении камеры, разработан метод объектной компенсации движения камеры на основе двухмерной объектной декомпозиции и алгоритм его реализации на основе монокулярной оптической модели. Суть метода заключается в сегментации и объектной декомпозиции опорного кадра видеоданных и предсказании очередного кадра с использованием информации об установке и перемещении видеокамеры. В отличие от известных методов эффективного кодирования видеоданных, использующих блочную сегментацию, отделение подвижных объектов от фона и вектора движения для каждого блока, разработанный метод использует объектно-ориентированную сегментацию, учитывающую структуру изображения, полную объектную декомпозицию, отделяющую неподвижные объекты друг от друга, и один вектор движения для предсказания. За счет этого обеспечивается уменьшение ошибки предсказания до 4,5 раз и увеличение коэффициента сжатия разности предсказанного и прогнозируемого кадров до 2,5 раз по сравнению с методом блочной компенсации движения, используемым в видеокодеке MPEG-4.

На основе предложенных методов и алгоритмов эффективного кодирования видеоданных разработаны программные модули видеокодека с объектной компенсацией движения камеры для сжатия видеоданных в мобильной системе и увеличения за счет этого ее пропускной способности. Результаты диссертационной работы использованы в КБ «Дисплей» и внедрены в учебный процесс в Белорусском государственном университете информатики и радиоэлектроники.

SUMMARY

Al-Juboori Tareq Mohammed Salman

EFFECTIVE VIDEO CODING IN MOBILE SYSTEMS BASED ON OBJECT MOTION COMPENSATION OF CAMERA

Keywords: efficient coding of video, object-based motion compensation cameras, video compression, monocular object decomposition, prediction of changes in video data.

The aim is to develop methods based on object-camera motion compensation for efficient encoding of video data in mobile systems.

The influence of the parallax at the coding efficiency of video data is generated by a moving camera. It is found that the use of coding which is based on video frame differencing and block motion compensation, with conditions of parallax has inefficient due to significant prediction error frames. For increase the compression ratio of video data which is generated by the motion of the camera, a method is developed by object motion compensation which is basis on two-dimensional camera object decomposition algorithm and its implementation on the basis of monocular optical model. The method consists of: the segmentation and object decomposing the reference frame of video data and make prediction of the next frame by using information about installation and movement of the camera. Unlike the known methods of efficient video coding using block segmentation, the moving objects at the background and motion vector for each block should be calculated; the method uses an object-oriented segmentation, based on the structure of the image, the full object decomposition between fixed objects from each other, and only one vector to predict the entire frame. This is to reduce the prediction error of up to 4.5 times and increased compression ratio difference between the predicted and predictable up to 2.5 times compared with the method of block motion compensation in codec MPEG-4 video compression.

On the foundation of the proposed methods and algorithms for efficient coding of video software modules developed video codec with object motion compensation for video compression for camera, in mobile systems and increase the bandwidth. The results of the thesis are used in Kb "Display" and introduced into the curriculum of the Belarusian State University of Informatics and Radio Electronics.

Научное издание

АЛЬ-ДЖУБУРИ Тарик Мохаммед Сальман

**ЭФФЕКТИВНОЕ КОДИРОВАНИЕ ВИДЕОДАНЫХ В МОБИЛЬНЫХ
СИСТЕМАХ НА ОСНОВЕ ОБЪЕКТНОЙ КОМПЕНСАЦИИ ДВИЖЕНИЯ
КАМЕРЫ**

Специальность 05.12.13 – Системы, сети
и устройства телекоммуникаций

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Подписано в печать 16.01.2012.	Формат 60x84 ¹ / ₁₆ .	Бумага офсетная.
Гарнитура «Таймс».	Отпечатано на ризографе.	Усл. печ. л. 1,63.
Уч.-изд. л. 1,4.	Тираж 60 экз.	Заказ 48.

Издатель и полиграфическое исполнение: учреждение образования
«Белорусский государственный университет информатики и радиоэлектроники»
ЛИ №02330/0494371 от 16.03.2009. ЛП №02330/0494175 от 03.04.2009
220013, Минск, П. Бровка, 6