

УДК 621.397

КОДИРОВАНИЕ С НЕРАВНОЙ ЗАЩИТОЙ КОМПРЕССИРОВАННОГО ВИДЕО В БЕСПРОВОДНЫХ СЕТЯХ

Ф. А. Таубин,

доктор техн. наук, профессор

А. В. Чуйков,

ассистент

Санкт-Петербургский государственный университет аэрокосмического приборостроения

Рассматривается схема передачи компрессированного видео, базирующаяся на разбиении битового потока в видеокодере на подпотоки (с учетом степени важности основных синтаксических единиц) и последующем помехоустойчивом кодировании с неравной защитой подпотоков от канальных ошибок. Для кодирования с неравной защитой используется класс совместимых по скорости перфорированных кодов повторения-накопления (RA-кодов), позволяющих обеспечить высокую помехоустойчивость и обладающих весьма привлекательными обменными соотношениями «помехоустойчивость — сложность реализации». Предлагается подход к согласованному выбору основных параметров рассматриваемой схемы передачи, включающий, в частности, обмен между скоростью компрессированного видео и скоростями RA-кодов.

Ключевые слова — потоковое видео, беспроводные сети, пиковое отношение сигнал/шум, помехоустойчивое кодирование с неравной защитой, коды повторения-накопления.

Введение

Передача потокового видео и обеспечение доступа к видеосервисам составляют доминирующую долю трафика в локальных и региональных беспроводных сетях [1]. Можно выделить несколько важных обстоятельств, связанных с реализацией доставки видеоконтента. Во-первых, в настоящее время в большинстве беспроводных сетей используются, как правило, протоколы передачи, не предусматривающие наличие QoS (Quality of Service) механизмов. Во-вторых, ограниченная полоса частот в канале диктует необходимость передачи видеоконтента в компрессированном виде. В-третьих, компрессированное видео весьма чувствительно к канальным ошибкам. Действительно, даже одиночная ошибка в компрессированном битовом потоке может привести к появлению последовательности, не восстанавливаемой видеодекодером, что в свою очередь может вызвать остановку декодирования. Кроме того, обратной стороной предикативного кодирования, используемого при компрессии, является пространственно-временное размножение ошибок. В результате обеспечение надежной доставки компрессирован-

ного видео (в отсутствие QoS-гарантии) оказывается весьма актуальной и перспективной задачей.

Стандартным инструментом обеспечения надежной передачи является использование помехоустойчивого кодирования. К настоящему времени предложено множество схем, демонстрирующих эффективность введения помехоустойчивого кодирования при пакетной передаче компрессированного видео (см. например, работы [2–6]). Среди предложенных схем следует выделить, как наиболее перспективные, методы передачи, основанные на использовании помехоустойчивого кодирования с неравной защитой [4–6]. Более высокая эффективность схем кодирования с неравной защитой достигается благодаря учету структуры компрессированного видео, и в частности — учету неравной значимости различных синтаксических элементов битового потока с точки зрения защиты от канальных ошибок.

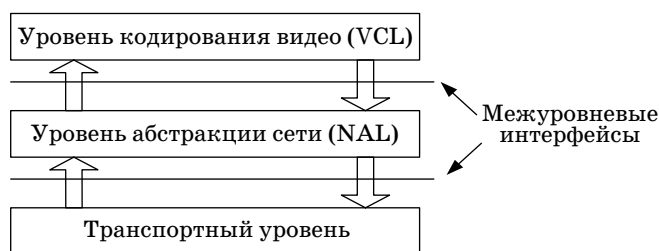
Как правило, при анализе схем с неравной защитой основное внимание уделялось кодовым конструкциям, основанным на совместимых по скорости сверточных кодах или турбокодах. В данной работе рассматривается метод передачи компрессированного видео с неравной защитой, базирующийся

ся на использовании совместимых по скорости кодов повторения-накопления, которые обладают рядом существенных достоинств по сравнению со сверточными кодами и турбокодами.

Структура компрессированного видео

Будем полагать, что компрессия видео осуществляется с использованием передового стандарта HEVC (High Efficiency Video Coding) [7], позволяющего существенно повысить степень компрессии по сравнению с наиболее популярным в настоящее время стандартом H.264/AVC. Компрессированное видео, порождаемое HEVC-кодеком, имеет сложную иерархическую организацию битового потока, которая в первую очередь определяется тем, что функциональные возможности кодера, как и в стандарте H.264/AVC, концептуально разделены на два уровня (рис. 1): уровень кодирования (компрессии) исходного видео VCL (Video Coding Layer) и уровень сетевой абстракции NAL (Network Abstraction Layer), обеспечивающий форматирование компрессированного видео для последующей передачи по сети.

На уровне VCL реализуется стандартная (как и в H.264/AVC) схема гибридного кодирования, включающая предсказание, двумерное косинусное преобразование и арифметическое кодирование. В большинстве приложений исходная видеопоследовательность разбивается на группы кадров GoP (Group of Pictures) и каждая группа кодируется независимо от других групп, что позволяет ограничить временное размножение ошибок. Для ограничения пространственного размножения ошибок каждый кадр сегментируется на части, называемые кодируемыми элементами. Кодируемый элемент представляет собой дерево, состоящее из квадратных блоков различного размера; при этом каждый кодируемый элемент может быть рекурсивным образом разбит с помощью четверичной сегментации на поддеревья меньшего размера. Другими словами, кодируемый элемент представляет совокупность блоков различного размера, связанных между собой дре-

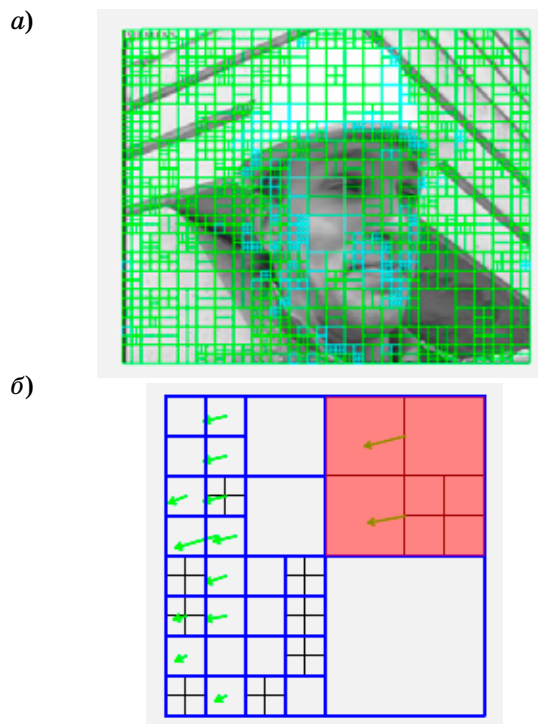


■ Рис. 1. Разделение функций в видеокодере

вовидной структурой (СТВ — Coding Tree Block) (рис. 2, а). В свою очередь, СТВ разбиваются на более мелкие единицы вплоть до так называемых единиц преобразования (TU — Transformation Unit) (рис. 2, б). Следует отметить, что новая терминология, используемая в HEVC для описания кодируемых элементов, пока не имеет общепринятого перевода на русский язык, поэтому для удобства изложения кодируемый элемент будем называть макроблоком (как в H.264/AVC).

В простейшем варианте битовый поток, соответствующий макроблоку, формируется как кодированная последовательность I- (интра) и P- (предсказанных) блоков. Каждый I-блок кодируется независимо, а при кодировании P-блока учитываются как предыдущие, так и последующие I- и (или) P-блоки. Очевидно, что с точки зрения защиты от канальных ошибок I-блоки являются более важными, чем P-блоки.

Уровень сетевой абстракции NAL разработан как дружественный протокол для хранения, передачи по сети и инкапсуляции с другими протоколами (например, RTP (Real-time Transport Protocol)). Протокол NAL форматирует компрессированное видео, определяя минимальный набор правил для надежного транспортирования потока, получаемого от VCL, по различным сетям.



■ Рис. 2. Иерархическое разбиение яркостной компоненты в HEVC: а — разбиение кадра на кодируемые элементы в виде СТВ; б — разбиение СТВ на TUs (с векторами компенсации)

Маркер синхронизации	Заголовок видеопакета	Данные СТВ макроблока
----------------------	-----------------------	-----------------------

■ Рис. 3. Структура видеопакета

Компрессированные видеоданные представляются в виде единиц NAL, называемых NALU (NAL Unit), каждая из которых является пакетом, состоящим из целого числа байт. Заголовок NALU, состоящий из первых двух байт, предназначен для описания типа данных. Остальные байты NALU представляют собой собственно видеoinформацию, определяемую заголовком. Единица NAL представляет единый формат как для байтового потока, так и для пакетной передачи.

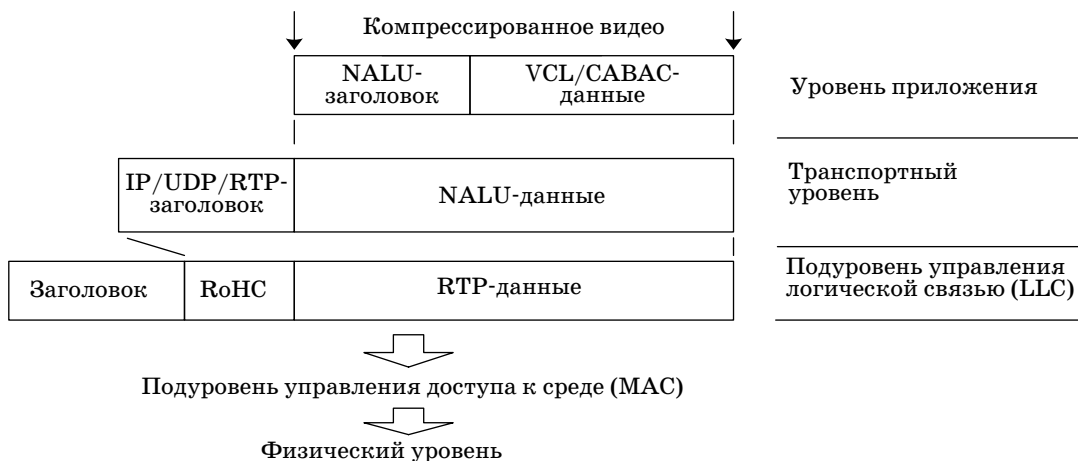
Форматированный поток представляет собой последовательность видеопакетов. Для потоковых транспортных протоколов, таких как H.320 или H.324M, при форматировании в состав видеопакета включается стартовый код (рис. 3), позволяющий транспортному уровню и декодеру легко идентифицировать структуру битового потока. При работе с пакетно-ориентированными протоколами, таки-

ми как RTP/IP и TCP/IP, NAL форматирует компрессированное видео без включения в видеопакеты стартовых кодов. В этом случае определение границ единиц NAL производится фрагментацией транспортного протокола системы.

С точки зрения защиты от канальных ошибок синтаксические элементы каждого макроблока, очевидно, имеют различную степень важности. Исходя из этого в HEVC-кодеке (как и в стандарте H.264/AVC) предусмотрена возможность разбиения содержимого блоков (Data Partition Mode) на три составных фрагмента (A, B и C) с учетом значимости данных, включаемых в каждый фрагмент. Полученные в результате разбиения три подпоток — A, B и C соответственно, кодируются отдельно, после чего инкапсулируются — каждый подпоток в свой собственный NAL-пакет, порождая соответственно NAL-A, NAL-B и NAL-C пакеты. В простейшем сценарии разбиения макроблоков пакет NAL-A содержит информацию о параметрах квантования и векторах компенсации движения, тогда как пакеты NAL-B и NAL-C несут информацию о текстуре для интраблоков и интерблоков соответственно (рис. 4).



■ Рис. 4. Простейший вариант разбиения макроблоков



■ Рис. 5. Схема инкапсуляции компрессированного видео

Общая схема инкапсуляции компрессированного видео применительно к передаче в беспроводной сети показана на рис. 5. Инкапсуляция осуществляется различными уровнями операционной системы путем добавления на каждом уровне необходимых заголовков и выполнения сегментации данных. Процедура RoHC (Robust Header Compression) осуществляет компрессию заголовков пакетов транспортного уровня.

Разбиение компрессированного битового потока на подпотоки

Включение опции Data Partition Mode позволяет организовать неравную защиту подпотоков А, В и С (с учетом степени их важности) от канальных ошибок и тем самым более эффективно использовать избыточность, вносимую при введении помехоустойчивого кодирования. Степень влияния канальных ошибок на отдельные синтаксические элементы компрессированного видео определяется, очевидно, структурой HEVC-кодека.

Эвристический анализ процесса функционирования HEVC-кодека при наличии канальных ошибок показывает, что декодер в результате ошибок в пакетах или потери фрагмента может вести себя в соответствии с одним из следующих сценариев:

1) декодировать компрессированное видео либо без ошибок (например, если ошибки/потери затронули данные, не участвующие в декодировании), либо с ошибками, порождающими локальные искажения восстановленного видео;

2) декодировать компрессированное видео с ошибками, порождающими существенные искажения восстановленного видео, включая полное искажение кадров;

3) остановить процесс декодирования.

Первый сценарий поведения декодера реализуется, как правило, при появлении ошибок в DC- и AC-компонентах двумерного косинусного преобразования блоков. Второй сценарий поведения декодера вызывают ошибки в векторах ком-

пенсации и заголовках макроблоков. К останову декодера приводят ошибки в заголовках RTP-пакетов и единиц NAL. Типичные искажения, характерные для первого и второго вариантов локализации ошибок, иллюстрируют рис. 6, а—г.

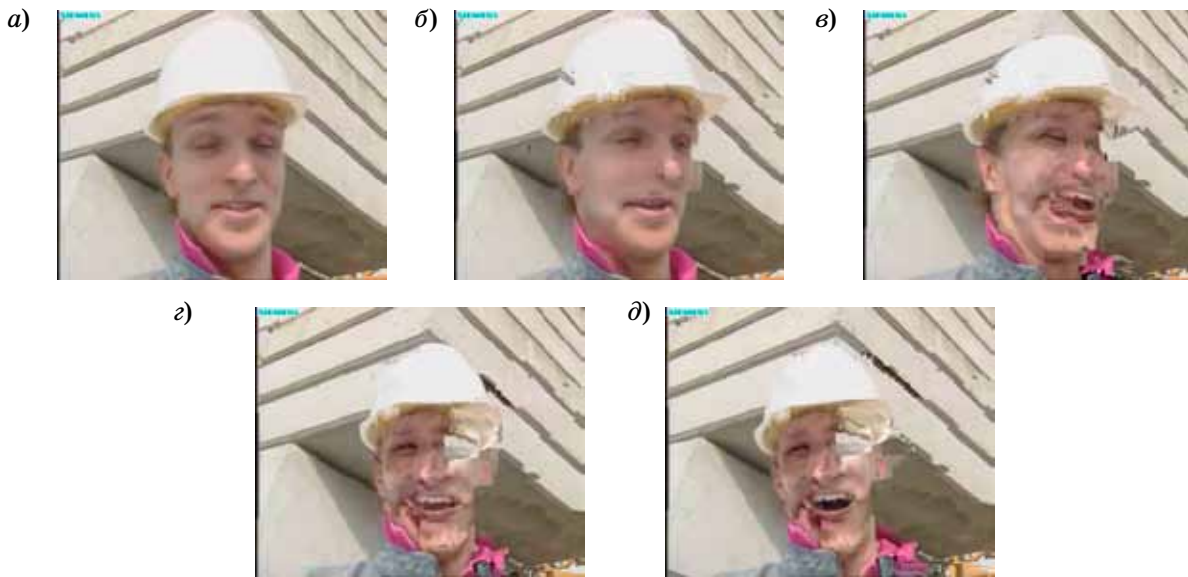
Следует отметить, что степень размножения искажений, вызываемых канальными ошибками, существенным образом зависит от количества кадров L , входящих в GoP. Дело в том, что канальные ошибки в начале GoP порождают более сильные искажения видео, чем ошибки в конце GoP, и степень искажений возрастает с увеличением размера GoP. Например, ошибка в первом Р-кадре после I-кадра приведет к искажению всех последующих кадров в GoP, тогда как канальная ошибка в последнем Р-кадре не приведет к искажению остальных кадров. Другими словами, начальные кадры GoP оказываются более значимыми (с точки зрения их влияния на степень размножения искажений), чем последующие, и степень дифференциации кадров GoP в этом смысле возрастает с увеличением размера GoP. Типичные искажения, вызываемые канальными ошибками в начале и в конце группы кадров GoP, показаны на рис. 7, а—д.

Характер локализации канальных ошибок внутри кадра также может влиять на степень порождаемых ими искажений. Так, если канальные ошибки поражают макроблок в начале кадра, то декодер теряет синхронизацию (вследствие использования кода с переменной длиной) и прекращает декодирование текущего видеокadra. Если же канальные ошибки попали на макроблоки, расположенные ближе к концу кадра, то подавляющую часть кадра декодер сможет успешно восстановить. Другими словами, макроблоки в начале кадра оказываются более значимыми, чем последующие макроблоки.

С учетом приведенных выше результатов эвристического анализа процесса функционирования HEVC-кодека представляется рациональным следующий вариант использования опции Data Partition Mode. Подпоток А наиболее значимых данных (биты класса А) включает:



■ Рис. 6. Иллюстрация искажений, порождаемых ошибками в векторах компенсации (а); заголовках макроблоков (б); в коэффициентах текстуры яркостных (в) и цветоразностных (г) компонент



■ **Рис. 7.** Иллюстрация искажений, порождаемых канальными ошибками:
a — в начале GoP; *b* — в кадре № 4 (первая ошибка); *c* — кадре № 12; *d* — кадре № 20; *e* — кадре № 28

- 1) размер изображения;
- 2) порядок следования кадров в GoP;
- 3) информацию о профиле;
- 4) глубину иерархии преобразования (Intra и Inter);
- 5) значения шага квантования.

Следующий по значимости данных подпоток В (биты класса В) включает:

- 1) результаты предсказания яркостных компонент;
- 2) коэффициенты двумерного косинусного преобразования для Intra и Inter (без знаков);
- 3) номер последнего значащего коэффициента в блоке (Intra).

Подпоток С наименее чувствительных более к канальным ошибкам данных (биты класса С) включает:

- 1) номер последнего значащего коэффициента в блоке (Inter);
- 2) векторы компенсации;
- 3) знаки коэффициентов двумерного косинусного преобразования для Intra и Inter;

- 4) направление предсказания цветоразностных компонент.

Отметим, что текстура I-кадров не включена в подпоток А по двум причинам. Во-первых, ее значимость, как показали наши эксперименты, все же заметно меньше, чем значимость заголовков RTP-пакетов и единиц NAL. Во-вторых, наличие канальных ошибок диктует необходимость ограничения количества кадров L в GoP сравнительно небольшими значениями, что в свою очередь нивелирует в известной степени разность в значимости I-кадра и последующих P-кадров. При сопоставлении искажений видео, вызываемых канальными ошибками в текстуре и векторах компенсации, оказалось, что в первом случае степень искажений заметно больше. Поэтому текстура включена в подпоток В, а векторы компенсации — в подпоток С. Для предложенного варианта разбиения компрессированного видеопотока на подпоток характерны а) весьма малая относительная доля α_A подпотока А и б) примерно сопоставимые по объему относительные

■ **Таблица 1.** Доли подпоточков А, В и С для клипа Foreman

Подпоток	Относительные доли подпоточков, %, при скорости компрессированного видео							
	$R=1500$ Кбит/с				$R=1700$ Кбит/с			
	$L=4$	$L=8$	$L=16$	$L=32$	$L=4$	$L=8$	$L=16$	$L=32$
α_A	1,29	0,75	0,36	0,16	1,20	0,73	0,36	0,35
α_B	51,36	39,74	44,07	52,73	53,31	42,16	44,36	40,94
α_C	47,35	59,51	55,57	47,11	45,49	57,11	55,28	58,71

■ Таблица 2. Доли подпотоков А, В и С при количестве кадров в GoP $L=16$

Скорость компрессированного видео R , Кбит/с	Относительные доли подпотоков, %		
	α_A	α_B	α_C
1000	0,47	40,58	58,95
1300	0,43	41,89	57,68
1500	0,36	44,07	55,57
1700	0,36	44,36	55,28

доли α_B и α_C подпотоков В и С соответственно. В табл. 1 приведены данные, иллюстрирующие типичные значения относительных долей подпотоков А, В и С для тестового клипа Foreman. Этот клип имеет разрешение CIF (352×288 пикселей, 8 бит на отсчет, формат YCbCr 4:2:0) и частоту кадров $f=30$ кадров/с.

Видно, что доля подпотока А уменьшается с увеличением количества кадров в GoP, тогда как соотношение α_B и α_C не имеет явно выраженной тенденции. Представляется правдоподобным, что ситуация, когда $\alpha_B < \alpha_C$, более предпочтительна при кодировании с неравной защитой в силу того, что подпоток С требует сравнительно меньшей защиты от канальных ошибок. В этом плане следует выделить вариант, соответствующий $L=16$ (табл. 2), для которого в широком диапазоне скоростей компрессированного видео условие $\alpha_B < \alpha_C$ выполняется с заметным перевесом доли подпотока С — отношение α_B/α_C находится в диапазоне от 1,25 до 1,45.

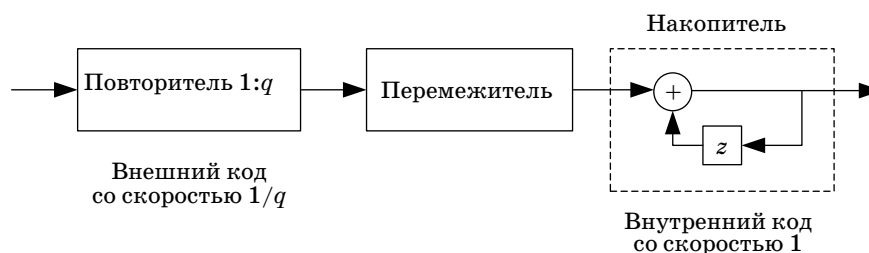
Помехоустойчивое кодирование подпотоков

Один из наиболее перспективных подходов к организации помехоустойчивого кодирования с неравной защитой основан на использовании совместимых по скорости перфорированных кодов [8–10]. Совокупность совместимых по скорости перфорированных кодов порождается низкоскоростным материнским кодом и конструируется в виде иерархии вложенных кодов, формируе-

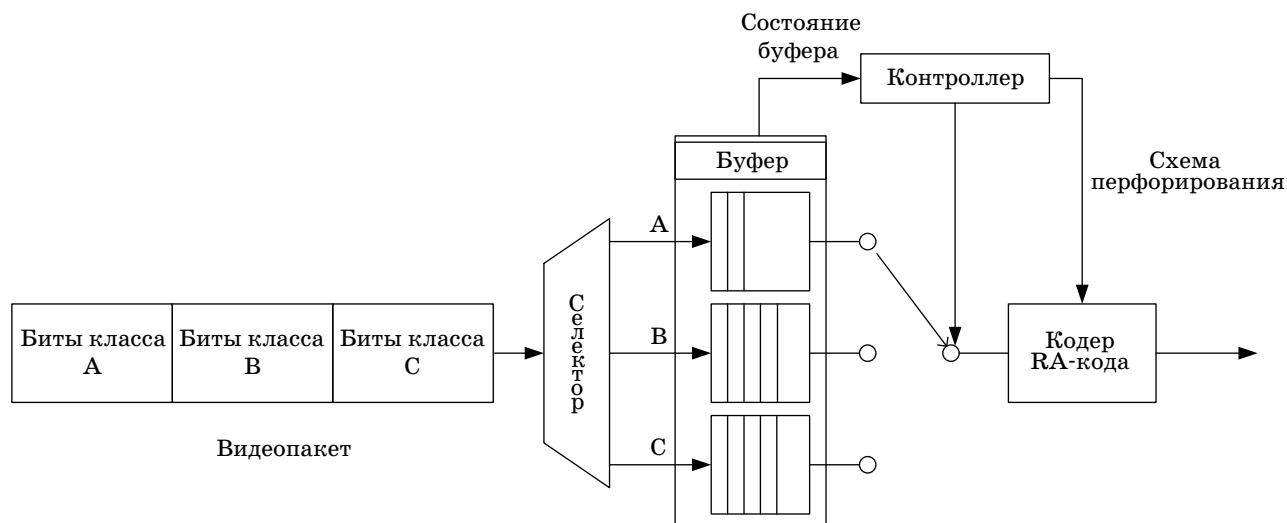
мой посредством последовательного перфорирования (выкалывания) части символов материнского кода. Коды в иерархии упорядочены по скоростям таким образом, что код с большей скоростью вложен в коды с меньшей скоростью. Другими словами, все кодовые биты кода с большей скоростью используются каждым кодом в этой иерархии, имеющим меньшую скорость. Важным достоинством совместимых по скорости перфорированных кодов является то, что как кодирование, так и декодирование всех кодов из этой совокупности выполняется кодером и декодером, имеющими универсальную структуру.

Среди предложенных к настоящему времени совместимых по скорости перфорированных кодов наиболее предпочтительным для организации неравной защиты подпотоков представляется класс кодов, в которых материнскими являются коды повторения-накопления (Repeat-Accumulate — RA). RA-коды [11] обеспечивают помехоустойчивость, сопоставимую с помехоустойчивостью, достигаемой при использовании как низкоплотностных, так и турбокодов, и при этом имеют линейную сложность кодирования и линейную сложность декодирования. Схема кодера RA-кода приведена на рис. 8.

RA-код представляет собой каскадную конструкцию, состоящую из двух весьма простых кодов, разделенных перемежителем; в этой конструкции внешним кодом является код с q -кратным повторением, а внутренним кодом (накопительным звеном) — рекурсивный сверточный код со скоростью 1 и передаточной характеристикой $1/(1+z)$. Таким образом, скорость материнского RA-кода равна $1/q$. Семейство совместимых по скорости кодов порождается из материнского RA-кода путем перфорирования а) информационных символов, б) символов на выходе внешнего кодера и в) символов на выходе внутреннего кодера. Схемы перфорирования (puncturing patterns) выбираются, как правило, на основе EXIT chart-анализа, позволяющего оптимизировать характеристики кодов. Кодер RA-кода может быть представлен с помощью графической модели —



■ Рис. 8. Кодер RA-кода



■ Рис. 9. Схема помехоустойчивого кодирования с неравной защитой

графа Таннера, что означает возможность декодирования RA-кодов с использованием итеративного алгоритма распространения доверия (Belief Propagation — BP), оперирующего с графом Таннера.

Схема помехоустойчивого кодирования с неравной защитой для компрессированного видео приведена на рис. 9. Битовый поток, представляющий собой последовательность видеопакетов, с выхода видеокодера поступает на вход селектора, где каждый видеопакет разбивается на фрагменты А, В и С в соответствии с изложенной выше процедурой реализации опции Data Partition Mode. Подпотoki, формируемые селектором, буферизуются и поочередно кодируются — каждый подпоток с использованием своего вложенного кода. Данные из буферов считываются блоками, имеющими длину k_A , k_B и k_C соответственно, при этом кодовые блоки для всех подпотоков имеют одну и ту же длину n . В приемнике декодированные блоки с выхода универсального RA-декодера распределяются между тремя буферами, после чего мультиплексируются в единый поток.

Обозначим через $r_A = k_A/n$, $r_B = k_B/n$ и $r_C = k_C/n$ скорости RA-кодов, используемых для помехоустойчивого кодирования подпотоков А, В и С соответственно. Тогда результирующая скорость битового потока, отправляемого в канал, составляет $R(\alpha_A/r_A + \alpha_B/r_B + \alpha_C/r_C)$.

Совместный выбор параметров схемы передачи с неравной защитой подпотоков

Качество передачи видео, как правило, оценивается пиковым отношением сигнал/шум (Peak

Signal-to-Noise Ratio — PSNR), определяемым как $PSNR = 10 \log_{10}(255/D)$ [дБ], где D — средняя (приходящаяся на один пиксель) дисперсия ошибки между исходным и принятым видео. Величина D весьма сложным образом зависит от параметров схемы передачи, включающей видеокодек, каналный кодек и беспроводную среду передачи. Поэтому во многих публикациях при исследовании качества передачи видео принимается ряд допущений, вполне обоснованных как эмпирически, так и посредством моделирования. Суть этих допущений сводится к использованию аддитивных моделей для учета искажений, возникающих как при компрессии, так и при передаче по каналу. Основное допущение состоит в том, что средняя дисперсия ошибки D может быть аппроксимирована суммой $D_S + D_C$, где D_S — средняя дисперсия ошибки, возникающая при компрессировании видео, а D_C — средняя дисперсия ошибки, порождаемая каналным шумом [12–14], так что оценивание средней дисперсии ошибки D сводится к оцениванию величин D_S и D_C . В работе [12] предложена модель, позволяющая связать дисперсию ошибки D_S и скорость компрессированного видео R в следующем виде:

$$D_S = D_0 + \theta(R - R_0), \quad (1)$$

где D_0 , R_0 и θ — параметры модели, определяемые свойствами компрессируемого видео, размером GoP и долей I-кадров. На рис. 10 представлен результат аппроксимации моделью (1) экспериментально вычисленных значений средней дисперсии ошибки D_S для тестового клипа Foreman.

Вычисление средней дисперсии ошибки, порождаемой каналным шумом, в общем случае

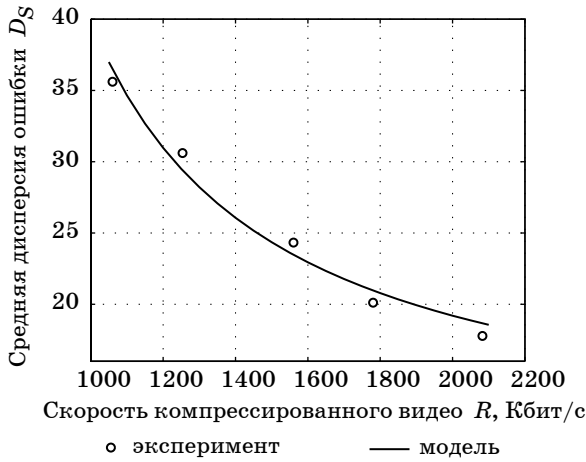


Рис. 10. Аппроксимация дисперсии ошибки D_S моделью (1) для клипа Foreman

представляет собой весьма сложную задачу. В качестве первого приближения для оценки величины D_C может быть применена предложенная в работе [12] техника оценивания, базирующаяся на использовании аддитивной модели для учета влияния канальных ошибок. Согласно этой технике, если канальная ошибка порождает в I-кадре искажение со средней дисперсией (на пиксель) σ_I^2 , то результирующая средняя дисперсия ошибки (с учетом размножения искажений внутри GoP) может быть оценена в виде

$$n_I P_e \sigma_I^2 \sum_{j=0}^{L-1} (1-j/L)/(1+\gamma_I j), \quad (2)$$

где n_I — среднее количество канальных ошибок, приходящихся на I-кадр; P_e — вероятность канальной ошибки; $\gamma_I, \gamma_I \in [0, 1]$ — параметр, зависящий от выбранного метода компенсации движения и схемы маскирования искажений. Использование подобного подхода для оценки величины D_C применительно к рассматриваемой модели передачи сводится к следующему.

1. Оценивание средней дисперсии ошибки (с учетом размножения искажений внутри GoP), порождаемой ошибкой в каждом из подпотоков. Обозначим эти величины как σ_A^2, σ_B^2 и σ_C^2 соответственно; отметим, что набор $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$ можно интерпретировать как своего рода чувствительность пары «клип-видеокодек» к канальным ошибкам в подпотоках.

2. Использование аддитивной модели для учета искажений, порождаемых ошибками на выходе канальных декодеров для подпотоков А, В и С. Полученные в результате оценки величин σ_A^2, σ_B^2 и σ_C^2 имеют следующий вид:

$$\sigma_A^2 \approx \sigma_{IA}^2 \beta \sum_{j=0}^{L-1} \frac{1-j/L}{1+\gamma_{IA} j} + \sigma_{PA}^2 \frac{1-\beta}{L-1} \sum_{l=1}^{L-1} \sum_{j=l}^{L-1} \frac{1-j/l}{1+\gamma_{PA} j}, \quad (3)$$

$$\sigma_B^2 \approx \sigma_{IB}^2 \beta \sum_{j=0}^{L-1} \frac{1-j/L}{1+\gamma_{IB} j} + \sigma_{PB}^2 \frac{1-\beta}{L-1} \sum_{l=1}^{L-1} \sum_{j=l}^{L-1} \frac{1-j/l}{1+\gamma_{PB} j}, \quad (4)$$

$$\sigma_C^2 \approx \sigma_{IC}^2 \beta \sum_{j=0}^{L-1} \frac{1-j/L}{1+\gamma_{IC} j} + \sigma_{PC}^2 \frac{1-\beta}{L-1} \sum_{l=1}^{L-1} \sum_{j=l}^{L-1} \frac{1-j/l}{1+\gamma_{PC} j}, \quad (5)$$

где β — относительный объем, занимаемый кодированными I-кадрами в общем битовом потоке; σ_{IA}^2 — средняя (на пиксель) дисперсия искажения в I-кадре, порождаемая ошибкой на выходе канального декодера подпотока А; σ_{PA}^2 — средняя (на пиксель) дисперсия искажения в P-кадре, порождаемая ошибкой на выходе канального декодера подпотока А (аналогичным образом определяются величины $\sigma_{IB}^2, \sigma_{PB}^2, \sigma_{IC}^2$ и σ_{PC}^2).

Оценка порождаемой канальным шумом средней дисперсии ошибки D_C , полученная в рамках аддитивной модели, имеет вид

$$D_C \approx \frac{RL}{f} [\alpha_A P_{eA} \sigma_A^2 + \alpha_B P_{eB} \sigma_B^2 + \alpha_C P_{eC} \sigma_C^2], \quad (6)$$

где f — частота кадров в видеопоследовательности; P_{eA}, P_{eB}, P_{eC} — вероятности ошибки на бит для канальных декодеров подпотоков А, В и С соответственно.

Для RA-кодов аналитические границы вероятности ошибки декодирования в настоящее время не известны. Вместе с тем при передаче по каналу с аддитивным белым гауссовым шумом (АБГШ) и значениях вероятности ошибки на бит P_e в диапазоне от 10^{-4} до 10^{-9} зависимость P_e от отношения сигнал/шум на символ h [дБ] можно достаточно точно аппроксимировать в виде

$$P_e \approx \exp_{10}[-u(r, n)h + v(r, n)], \quad (7)$$

где коэффициенты $u(r, n)$ и $v(r, n)$ определяются скоростью r и длиной кодового блока n RA-кода. Например, для рассматриваемого в работе материнского RA-кода со скоростью $r=1/2$ и длиной кодового блока $n=1024$ коэффициенты $u(r, n)$ и $v(r, n)$ принимают значения 5,1 и -9,6 соответственно.

Комбинируя выражения (1), (6) и (7), получаем следующую оценку для средней дисперсии ошибки:

$$D \approx D_0 + \theta/(R - R_0) + \frac{RL}{f} [\alpha_A P_{eA} \sigma_A^2 + \alpha_B P_{eB} \sigma_B^2 + \alpha_C P_{eC} \sigma_C^2], \quad (8)$$

где

$$Pe_A \approx \exp_{10}[-u(r_A, n)h + v(r_A, n)],$$

$$Pe_B \approx \exp_{10}[-u(r_B, n)h + v(r_B, n)],$$

$$Pe_C \approx \exp_{10}[-u(r_C, n)h + v(r_C, n)],$$

а величины σ_A^2 , σ_B^2 и σ_C^2 оцениваются правыми частями в (3), (4) и (5) соответственно.

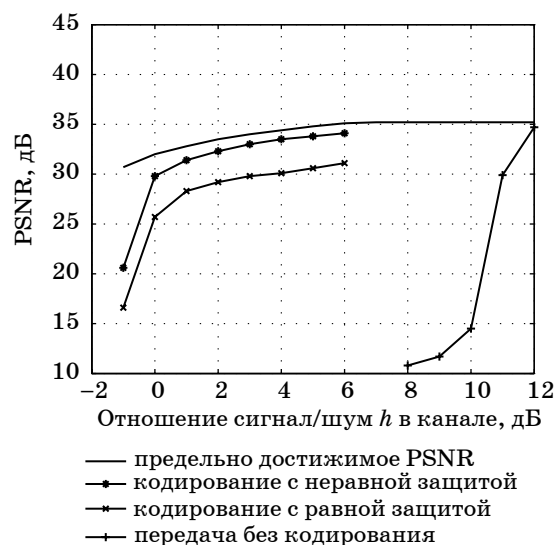
Наличие оценки (8) позволяет сформулировать задачу согласованного выбора основных параметров схемы передачи компрессированного видео по каналу с АБГШ. Под основными параметрами будем понимать размер GoP L , скорость компрессированного видео R и скорости r_A , r_B и r_C кодов, используемых для кодирования подпотоков; обозначим набор этих параметров как $S = (L, R, r_A, r_B, r_C)$. Согласованный выбор параметров подразумевает такой выбор набора S , который для заданного отношения сигнал/шум h в канале минимизирует правую часть (8) при выполнении следующих двух условий. Во-первых, результирующая скорость передачи не должна превышать заданной скорости передачи V в канале, т. е. $R(\alpha_A/r_A + \alpha_B/r_B + \alpha_C/r_C) \leq V$. Во-вторых, длина кодового блока n используемого материнского RA-кода должна быть ограничена сверху величиной, определяемой, исходя из допустимых значений задержки и сложности декодирования. Отметим, что представленный подход к согласованному выбору параметров может быть распространен, с использованием результатов работы [15], на топологии с промежуточными и interfering узлами.

Оценка эффективности введения неравной защиты подпотоков

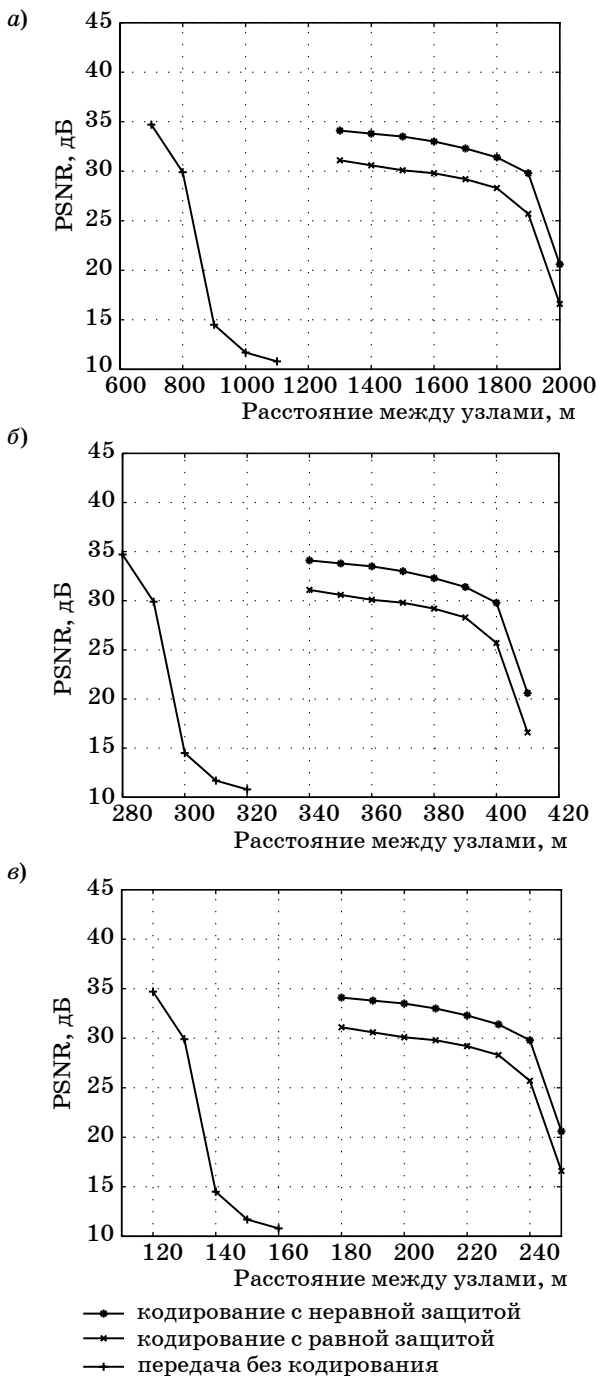
Сложность выражений, связывающих PSNR в доставленном видео и параметры схемы передачи с неравной защитой, препятствует получению достаточно наглядных аналитических оценок эффективности использования такой схемы применительно к произвольному видео (или некоторому классу видеопоследовательностей). В этой ситуации разумным компромиссом представляется получение численных оценок эффективности применительно к конкретным тестовым видеоклипам. В качестве примера рассмотрим согласованный выбор параметров в рамках обсуждаемой схемы при передаче тестового клипа Foreman между двумя узлами Wi-Fi сети со скоростью $V = 2$ Мбит/с. Будем полагать, что interfering и промежуточные узлы отсутствуют, что позволяет использовать модель канала с АБГШ. Допустимая длина кодового блока $n = 1024$, а скорости r_A , r_B и r_C кодов, используемых для кодирования подпотоков, выбираются из множества $\{s/(s+1) | 1 \leq s \leq 9\}$. Значения PSNR

в доставленном видео, соответствующие согласованному выбору параметров, приведены на рис. 11. Для сравнения на этом же рисунке приведены значения PSNR, получаемые а) при кодировании с равной защитой, б) при передаче без введения кодирования и в) при передаче со скоростью, сколь угодно близкой к пропускной способности канала (предельно достижимые значения PSNR). Представленные на рис. 11 результаты позволяют сделать следующие выводы. Во-первых, при передаче без кодирования удается обеспечить приемлемое качество доставленного видео лишь при отношении сигнал/шум в канале $h > 10,5$ дБ, тогда как при использовании RA-кодов этот порог снижается до $-0,5 \dots 0$ дБ. Во-вторых, кодирование с неравной защитой обеспечивает выигрыш примерно в 4 дБ относительно кодирования с равной защитой. В-третьих, при отношении сигнал/шум в канале $h \geq 0$ дБ схема кодирования с неравной защитой обеспечивает качество, достаточно близкое к предельно возможному — разница составляет от 1,1 до 2,3 дБ.

С прикладной точки зрения при передаче видео в беспроводной сети более наглядными (по сравнению с графиками на рис. 11) могут оказаться характеристики, иллюстрирующие качество принятого видео в зависимости от расстояния между передающим и приемными узлами. Соотношение между мощностью принятого сигнала и расстоянием между узлами определяется моделью, характеризующей степень ослабления сигнала при его распространении в пространстве. В работе далее рассматриваются три популярные



■ Рис. 11. Значения PSNR в доставленном видео в зависимости от отношения сигнал/шум в канале (клип Foreman, скорость передачи $V = 2$ Мбит/с)



■ **Рис. 12.** Значения PSNR в доставленном видео в зависимости от расстояния для различных моделей затухания (клип Foreman, скорость передачи $V=2$ Мбит/с): *a* — с квадратичным затуханием; *b* — с двухлучевым распространением сигнала; *в* — для экспоненциальной модели

модели [16]. Модель #1 — это простейшая модель с квадратичным затуханием, которая характеризует ослабление сигнала в открытом пространстве. Модель #2 предполагает двухлучевое распространение сигнала: один луч — прямой, вто-

рой — отраженный от поверхности Земли. Модель #3 определяет степень затухания в виде экспоненциальной функции от расстояния, причем показатель экспоненты может принимать три различных значения (в зависимости от расстояния). Для всех рассмотренных моделей полагалось, что несущая частота равна 2,412 ГГц, мощность передатчика имеет типовое значение 10 дБм, коэффициенты усиления передающей и приемной антенн равны 1, а уровень спектральной плотности мощности (теплового) шума составляет -174 дБм на 1 Гц полосы, что соответствует мощности шума -101 дБм в стандартной полосе 20 МГц.

Графики, иллюстрирующие зависимость PSNR в доставленном видео от расстояния между узлами для трех моделей распространения сигнала, представлены на рис. 12, *a–в*. Как видим, введение предложенной схемы кодирования с неравной защитой позволяет существенно увеличить максимально возможное расстояние между узлами. При этом следует отметить два важных обстоятельства. Во-первых, при требуемом значении PSNR в восстановленном видео больше 30 дБ (отличное качество по шкале MOS) выигрыш в максимальной дальности (относительно некодированной передачи) быстро снижается, тогда как выигрыш в максимальной дальности относительно кодирования с равной защитой возрастает. Во-вторых, для моделей #2 и #3 (с более сильным ослаблением передаваемого сигнала) выигрыш в максимальной дальности заметно меньше, чем для модели #1. Так, при требуемом значении PSNR = 30 дБ максимальная дальность для модели #1 возрастает более чем на 1000 м по сравнению с вариантом некодированной передачи и примерно на 400 м по сравнению со схемой кодирования с равной защитой, тогда как для моделей #2 и #3 этот выигрыш составляет примерно 110 и 40 м соответственно.

Заключение

В работе рассматривалась схема передачи компрессированного видео, базирующаяся на разбиении битового потока в видеокодере на три подпотока (с учетом степени важности основных синтаксических единиц) и последующем помехоустойчивом кодировании с неравной защитой подпотоков от канальных ошибок. Для организации кодирования с неравной защитой был выбран класс совместимых по скорости перфорированных кодов повторения-накопления (RA-кодов), обеспечивающих помехоустойчивость, сопоставимую с помехоустойчивостью низкоплотностных и турбокодов, и имеющих при этом линейную сложность кодирования-декодирования. Введение неравной за-

щиты, должным образом согласованной а) со скоростью компрессированного видео, б) со степенью чувствительности доставленного получателю видео к ошибкам при передаче и в) с отношением сигнал/шум в канале, позволило более эффективно использовать избыточность, вносимую при введении помехоустойчивого кодирования. Пример с передачей стандартного тестового клипа Foreman показывает, что кодирование с неравной защитой обеспечивает выигрыш по PSNR (в доставленном видео) примерно 4 дБ относительно кодирования с равной защитой. Важно также отметить, что рассматриваемая схема кодирования с неравной защитой обеспечивает качество, достаточно близ-

кое к предельно возможному — разница составляет от 1,1 до 2,3 дБ при отношении сигнал/шум в канале $h \geq 0$ дБ. При фиксированном значении PSNR в доставленном видео эффективность предложенной схемы кодирования с неравной защитой может быть оценена степенью увеличения максимально возможного расстояния между передающим и приемным узлами. Так, для трех рассмотренных в работе моделей ослабления сигнала в канале при требуемом значении PSNR=30 дБ возрастание максимальной дальности составляет от 38 до 137 % по сравнению с вариантом некодированной передачи и от 11 до 27 % по сравнению со схемой кодирования с равной защитой.

Литература

1. Cai L., Shen X., Mark J. W. *Multimedia Services in Wireless Internet: Modeling and Analysis*. – Wiley, 2009. – 290 p.
2. Nafaa A., Taleb T., Murphy L. Forward Error Correction Strategies for Media Streaming over Wireless Networks // *IEEE Communication Magazine*. Dec. 2007. P. 72–79.
3. Xiao J., Tillo T., Lin C., Zhao Y. Real-Time Forward Error Correction for Video Transmission // *Proc. of IEEE Conf. on Visual Communications and Image Processing VCIP'11*, Nov. 2011. P. 1–4.
4. Wang X., Tang K., Cui H. A Novel Unequal Error Protection Scheme for ROI based Video Coding in H.264/AVC // *Proc. of the Second Intern. Conf. on Computer Modeling and Simulation ICCMS'10*, Jan. 2010. P. 175–178.
5. Lamy-Bergot C., Chaunru N., Bergeron C. Unequal Error Protection for H.263+Bitstreams over a Wireless IP Networks // *Proc. of the IEEE ICASSP'06 Conf.*, Toulouse, May 2006. P. 377–380.
6. Stockhammer T., Bystrom M. H.264/AVC Data Partitioning for Mobile Video Communication // *Proc. of the Intern. Conf. on Image Processing*, Singapore, 2004. P. 545–548.
7. Sullivan G. J. et al. Overview of the High Efficiency Video Coding (HEVC) Standard // *IEEE Transactions on Circuits and Systems for Video Technology*. Dec. 2012. Vol. 22. N 12. P. 1649–1668.
8. Hagenauer J. Rate-Compatible Punctured Convolutional Codes (RCPC Codes) and their Applications // *IEEE Transactions on Communications*. Apr. 1988. Vol. 36. N 4. P. 389–400.
9. Barbulescu A. S., Pietrobon S. S. Rate-Compatible Turbo Codes // *Electronics Letters*. Mar. 1995. Vol. 31. N 7. P. 535–536.
10. Kim J., Ramamoorthy A., McLaughlin S. W. The Design of Efficiently-Encodable Rate-Compatible LDPC Codes // *IEEE Transactions on Communications*. Feb. 2009. Vol. 57. N 2. P. 365–375.
11. Planjery S., Gulliver T. A., Thangaraj A. Rate-Compatible Punctured Systematic Repeat-Accumulate Codes // *Proc. of IEEE Wireless Communications and Networking Conf.*, Mar. 2007. P. 718–722.
12. Stuhlmuller K., Farber N., Link M., Girod B. Analysis of Video Transmission over Lossy Channels // *IEEE J. on Selected Areas in Communications*. June 2000. Vol. 18. N 6. P. 1012–1032.
13. He Z. H., Cai J. F., Chen C. W. Joint Source Channel Rate-Distortion Analysis for Adaptive Mode Selection and Rate Control in Wireless Video Coding // *IEEE Transactions on Circuits and Systems for Video Technology*. June 2002. Vol. 12. P. 511–523.
14. Dani J. U., He Z. H., Xiong Z. Transmission Distortion Modeling for Wireless Video Communications // *Proc. of IEEE Global Telecommunications Conf. GLOBECOM'05*, Dec. 2005. P. 2425–2428.
15. Таубин Ф. А., Чуйков А. В. Анализ качества передачи потокового видео в беспроводных AD-НОС-сетях // *Информационно-управляющие системы*. 2012. № 2. С. 39–47.
16. Erceg V. et al. An Empirically Based Path Loss Model for Wireless Channels in Suburban Environments // *IEEE J. on Selected Areas in Communications*. July 1999. Vol. 17. N 7. P. 1205–1211.