



МЕТОДОЛОГИЧЕСКАЯ БАЗА УПРАВЛЕНИЯ БИТОВОЙ СКОРОСТЬЮ ПРИ ФОРМИРОВАНИИ ПРЕДСКАЗАННЫХ КАДРОВ

БАРАННИК В.В., ХАРЧЕНКО Н.А.,
БЕКIROV А.Э.

Строится методологическая база для управления битовой скоростью сжатой видеoinформации, что необходимо для ее адаптации относительно пропускной способности канала в телекоммуникационной сети. Разрабатываются выражения, позволяющие оценить объем передаваемых данных сжатого видеопотока для предсказанных кадров при различных вариантах выбора типа формируемых блоков. Учитываются следующие параметры: особенность цветовой модели, формат представления цветового пространства, тип кодирования видеоданных, размер блока и макроблока в кадре.

Введение

Передача цифрового видео от источника (видеокамера или записанный видеоролик) к получателю (видео-дисплей) вовлекает в работу целую цепь различных компонентов и процессов. Ключевыми звеньями этой цепи являются процессы компрессии (кодирования) и декомпрессии (декодирования). Здесь изначальный частотоемкий цифровой видеосигнал сокращается до размеров, подходящих для его дальнейшей передачи или хранения. Новые улучшенные разработки процессов компрессии и декомпрессии могут дать существенное коммерческое и техническое преимущество продукта. При этом обеспечивается лучшее качество видеоизображения, большая надежность и/или гибкая приспособляемость по сравнению с уже существующими решениями. Однако наряду с параметрами сжатия необходимо также вводить контроль битовой скорости видеопотока, который поступает в канал связи. Это обусловлено тем, что при обработке разных типов кадров формируется неравномерное количество бит на кадр. Это приводит к возникновению пульсации трафика, что негативно влияет на функционирование сети передачи данных. Поэтому есть необходимость в разработке алгоритма управления битовой скоростью видеопотока и его согласования с характеристиками ТКС. В работах [3,4] создан алгоритм обработки Р-кадров, который основан на использовании двух типов сегментов, выбираемых по заданному пределу с учетом функционала, оценивающего информативность полиадического кодирования на заключительном этапе сжатия видеоданных. По сравнению со стандарт-

ными методами кодирования он является предпочтительным. Однако для него отсутствует методологическая база оценки битовой скорости видеопотока в расчете на один кадр, разработка которой и является целью данной работы.

Основная часть

В [4] приведен алгоритм обработки блоков яркостной и цветоразностных составляющих представления предсказанного кадра. Определены длины кодовой последовательности передаваемых данных после сжатия для одного блока. Для яркостной составляющей блока $V(t, Y)_{k,l}^{(\xi, \gamma)}$ она определяется выражением:

$$d(t, Y)_{k,l}^{(\xi, \gamma)} = \begin{cases} d(P, Y)_{k,l}^{(\xi, \gamma)} = d(P, Y)_{t, \text{inf}}^{(\xi, \gamma)} + \\ + d(P, Y)_{t,s}^{(\xi, \gamma)}, \rightarrow f(E(t, Y)_{k,l}^{(\xi, \gamma)}) \leq \Delta D; \\ d(I, Y)_{k,l}^{(\xi, \gamma)} = d(I, Y)_{t, \text{inf}}^{(\xi, \gamma)} + \\ + d(I, Y)_{t,s}^{(\xi, \gamma)}, \rightarrow f(E(t, Y)_{k,l}^{(\xi, \gamma)}) > \Delta D, \end{cases} \quad (1)$$

где $d(P, Y)_{k,l}^{(\xi, \gamma)}$ – длина кодовой последовательности сжатых данных для Р-блока; $d(I, Y)_{k,l}^{(\xi, \gamma)}$ – длина кодовой последовательности сжатых данных для I-блока; $d(I, Y)_{t,s}^{(\xi, \gamma)}$ и $d(P, Y)_{t,s}^{(\xi, \gamma)}$ – служебные части для блоков типа I и P; $d(I, Y)_{t, \text{inf}}^{(\xi, \gamma)}$ и $d(P, Y)_{t, \text{inf}}^{(\xi, \gamma)}$ – длины кодового представления всех столбцов блоков типа I и P; (ξ, γ) – координаты макроблока в кадре, $\xi = \overline{1, m_{\text{mc}}}$; $\gamma = \overline{1, n_{\text{mc}}}$, здесь m_{mc} и n_{mc} – количество макроблоков в кадре по горизонтали и вертикали; (k, l) – координаты блока в макроблоке, $k = \overline{1, 2}$; $l = \overline{1, 2}$; t – номер текущего обрабатываемого кадра.

Процесс формирования кодовой последовательности для блока $V(t, Y)_{k,l}^{(\xi, \gamma)}$ показан на рис. 1.

Соответственно количество разрядов для блоков $V(t, C_r)_{k,l}^{(\xi, \gamma)}$ и $V(t, C_b)_{k,l}^{(\xi, \gamma)}$ после полиадического кодирования определяется следующими выражениями:

$$\begin{aligned} d(P, C_r)_{k,l}^{(\xi, \gamma)} &= d(P, C_r)_{t,s}^{(\xi, \gamma)} + d(P, C_r)_{t, \text{inf}}^{(\xi, \gamma)}, \\ d(P, C_b)_{k,l}^{(\xi, \gamma)} &= d(P, C_b)_{t,s}^{(\xi, \gamma)} + d(P, C_b)_{t, \text{inf}}^{(\xi, \gamma)}, \end{aligned} \quad (2)$$

где $d(P, C_r)_{t, \text{inf}}^{(\xi, \gamma)}$, $d(P, C_b)_{t, \text{inf}}^{(\xi, \gamma)}$ – длины информационной части соответственно для хроматической красной и хроматической синей составляющей; $d(P, C_r)_{t,s}^{(\xi, \gamma)}$, $d(P, C_b)_{t,s}^{(\xi, \gamma)}$ – длина служебной части соответственно для хроматической красной и хроматической синей составляющей.

Процесс формирования кодовой последовательности для блоков $V(t, C_r)_{k,l}^{(\xi, \gamma)}$ и $V(t, C_b)_{k,l}^{(\xi, \gamma)}$ показан на рис. 2.

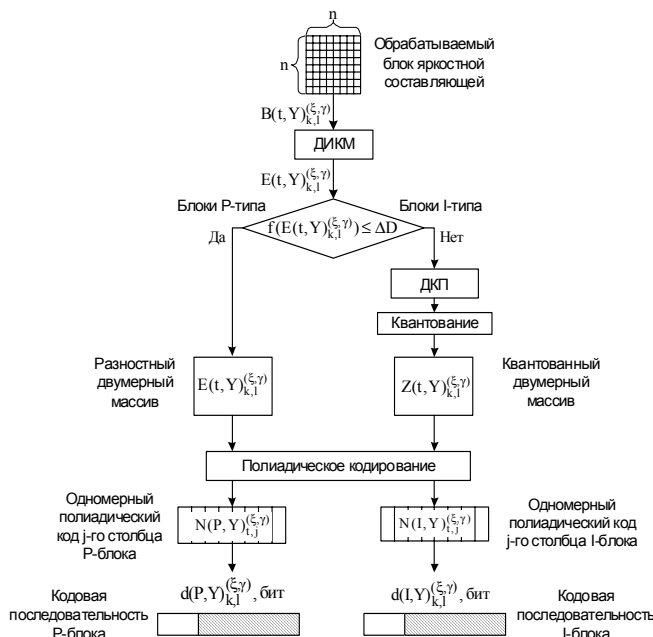


Рис. 1. Преобразование блока яркостной составляющей в кодovou последовательность

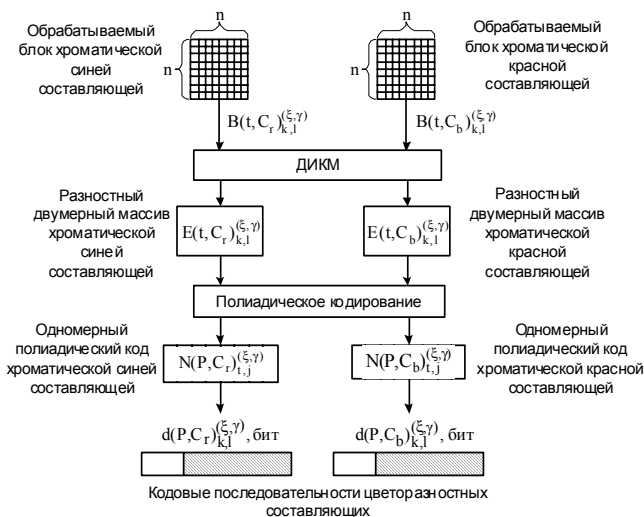


Рис. 2. Преобразование блоков цветоразностных составляющих в кодovou последовательность

При определении общего объема передаваемых данных нужно также учитывать цифровой формат макроблока видеоданных. Применяют три основных формата видео в зависимости от его назначения: 4:4:4, 4:2:2 и 4:2:0.

Как было описано ранее, макроблок после преобразования цветового пространства представляет собой совокупность сигнала яркости Y и двух цветоразностных сигналов C_r и C_b . Вариации их значений допускают 256 градаций (от 0 до 255 для Y и от -128 до 127 для C_r/C_b), что в двоичном исчислении соответствует 8 битам или 1 байту. Теоретически каждый элемент кадра имеет собственные значения $Y_{C_r C_b}$, т.е. требует 3 байта. Такое представление, когда как яркость, так и сигналы цветности имеют равное число независимых значений, обычно обозначают как 4:4:4 (рис. 3).

С учетом меньшей чувствительности зрительной системы человека к цветоразностным компонентам без видимой потери качества число цветовых отсчетов в каждой строке можно уменьшить вдвое.

Именно такое представление, обозначаемое как 4:2:2, было принято в вещательном телевидении. При этом для передачи полного значения телевизионного сигнала в каждом отсчете кадра достаточно 2 байта (чередуя через отсчет независимые значения C_r и C_b).

Более того, для целей потребительского видео признано допустимым уменьшение вдвое и вертикального цветового разрешения, т.е. переход к представлению 4:2:0. Это уменьшает приведенное число байт на отсчет до 1,5. Именно такое представление было заложено в DV-формат цифровых камер, а также формат DVD-видео.

Для учета дифференцированного влияния цифрового формата на объем битового описания кадра введем понятие структурной единицы кадра.

Определение. Структурной единицей кадра будем называть такую структуру, которая формируется на основе макроблоков по одному от каждой составляющей цветовой модели. Структурная единица $S(t)_{\text{пoo}}^{(\xi, \gamma)}$ задается следующей формулой:

$$S(t)_{\text{пoo}}^{(\xi, \gamma)} = S(t, Y)_{k,l}^{(\xi, \gamma)} \cup S(t, C_r)_{k,l}^{(\xi, \gamma)} \cup S(t, C_b)_{k,l}^{(\xi, \gamma)}, \quad (3)$$

где $S(t, Y)_{k,l}^{(\xi, \gamma)}$ – макроблок яркостной составляющей для текущего Р-кадра; $S(t, C_r)_{k,l}^{(\xi, \gamma)}$ и $S(t, C_b)_{k,l}^{(\xi, \gamma)}$ – цветоразностные составляющие, представленные соответственно хроматической красной и хроматической синей; (ξ, γ) – координаты макроблока в кадре, $\xi = \overline{1, m_{\text{mc}}}$; $\gamma = \overline{1, n_{\text{mc}}}$, здесь m_{mc} и n_{mc} – количество макроблоков в кадре по горизонтали и вертикали; t – номер текущего обрабатываемого кадра.

Определим количество бит, затраченное на представление одного макроблока, которое равно суммарному количеству бит на представление входящих в него блоков:

$$d(t, Y)_{k,l}^{(\xi, \gamma)} = \sum_{k=1}^{m_c} \sum_{l=1}^{n_c} d(t, Y)_{k,l}^{(\xi, \gamma)};$$

$$d(t, C_r)_{k,l}^{(\xi, \gamma)} = \sum_{k=1}^{m_c} \sum_{l=1}^{n_c} d(P, C_r)_{k,l}^{(\xi, \gamma)};$$

$$d(t, C_b)_{k,l}^{(\xi, \gamma)} = \sum_{k=1}^{m_c} \sum_{l=1}^{n_c} d(P, C_b)_{k,l}^{(\xi, \gamma)}, \quad (4)$$

где $d(t, Y)_{k,l}^{(\xi, \gamma)}$, $d(P, C_r)_{k,l}^{(\xi, \gamma)}$ и $d(P, C_b)_{k,l}^{(\xi, \gamma)}$ – соответственно объемы в битах для блоков $V(t, Y)_{k,l}^{(\xi, \gamma)}$, $V(t, C_r)_{k,l}^{(\xi, \gamma)}$ и $V(t, C_b)_{k,l}^{(\xi, \gamma)}$ для (ξ, γ) -х макроблоков яркостной и цветоразностных составляющих Р-кадра; m_c – количество сегментов в макросегменте по горизонтали; n_c – количество сегментов в макросегменте по вертикали.

Тогда общие объемы составляющих цветовой модели для одного кадра представляют собой суммарные объемы цифрового описания макроблоков $d(t, Y)^{(\xi, \gamma)}$, $d(t, C_r)^{(\xi, \gamma)}$, $d(t, C_b)^{(\xi, \gamma)}$. В соответствии с этим битовый объем составляющих цветовой модели будет определяться следующими выражениями:

$$\begin{aligned}
 D(t, Y) &= \sum_{\xi=1}^{m_{mc}} \sum_{\gamma=1}^{n_{mc}} d(t, Y)^{(\xi, \gamma)} ; \\
 D(t, C_r) &= \sum_{\xi=1}^{m_{mc}} \sum_{\gamma=1}^{n_{mc}} d(t, C_r)^{(\xi, \gamma)} ; \\
 D(t, C_b) &= \sum_{\xi=1}^{m_{mc}} \sum_{\gamma=1}^{n_{mc}} d(t, C_b)^{(\xi, \gamma)} , \quad (5)
 \end{aligned}$$

где m_{mc} и n_{mc} – количество макроблоков в кадре по горизонтали и вертикали.

Суммарная длина $D(t)$ кодовой последовательности цифрового описания составляющих цветовой модели на один P-кадр определяется выражением:

$$D(t) = D(t, Y) + D(t, C_r) + D(t, C_b) , \quad (6)$$

здесь $D(t, Y)$ – количество бит на представление яркостных составляющих цветовой модели; $D(t, C_r)$, $D(t, C_b)$ – количество бит на представление хроматических составляющих кадра.

С учетом соотношений (5) выражение (6) для величины $D(t)$ примет вид

$$\begin{aligned}
 D(t) &= \sum_{\xi=1}^{m_{mc}} \sum_{\gamma=1}^{n_{mc}} d(t, Y)^{(\xi, \gamma)} + \sum_{\xi=1}^{m_{mc}} \sum_{\gamma=1}^{n_{mc}} d(t, C_r)^{(\xi, \gamma)} + \\
 &+ \sum_{\xi=1}^{m_{mc}} \sum_{\gamma=1}^{n_{mc}} d(t, C_b)^{(\xi, \gamma)} . \quad (7)
 \end{aligned}$$

Из рис. 3 видно, что объемы составляющих кадра определяются форматом представления цветного пространства. Значит, объем цифрового описания $d(t)_{\text{код}}^{(\xi, \gamma)}$ для структурной единицы $S(t)_{\text{нод}}^{(\xi, \gamma)}$ зависит от объемов входящих в нее макроблоков и соответственно объемов цифрового описания входящих в структурную единицу блоков:

$$d(t)_{\text{код}}^{(\xi, \gamma)} = \sum_{k=1}^{m_c} \sum_{l=1}^{n_c} \left(d(t, Y)_{k,l}^{(\xi, \gamma)} + d(P, C_r)_{k,l}^{(\xi, \gamma)} + d(P, C_b)_{k,l}^{(\xi, \gamma)} \right) . \quad (8)$$

Тогда объем передаваемых видеоданных в расчете на один кадр $D(t)$ определяется затратами количества бит на представление всех структурных единиц $d(t)_{\text{код}}^{(\xi, \gamma)}$ кадра с учетом форматов представления цветного пространства.

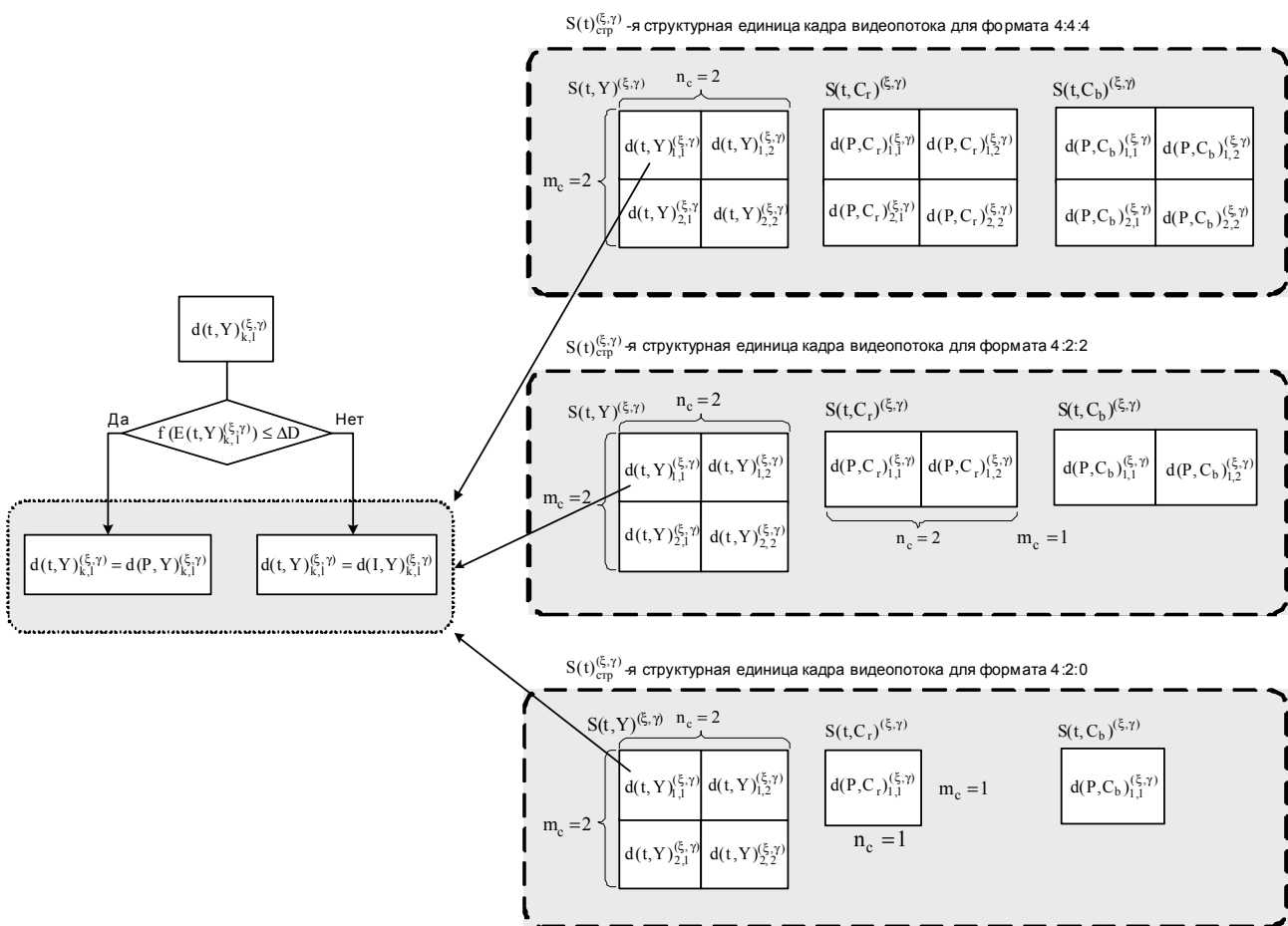


Рис. 3. Формирование структурной единицы кадра видеопотока с учетом цифровых форматов видео

Выражения (1), (2) для определения объема видеопотока в расчете на один кадр, с учетом выражений (7), (8), принимает вид:

$$D(t) = \sum_{\xi=1}^{m_{mc}} \sum_{\gamma=1}^{n_{mc}} \left(\begin{array}{l} \sum_{k=1}^{m_c} \sum_{l=1}^{n_c} d(t, Y)_{k,l}^{(\xi, \gamma)} + \\ + \sum_{k=1}^{m_c} \sum_{l=1}^{n_c} d(P, C_r)_{k,l}^{(\xi, \gamma)} + \\ + \sum_{k=1}^{m_c} \sum_{l=1}^{n_c} d(P, C_b)_{k,l}^{(\xi, \gamma)} \end{array} \right), \quad (9)$$

Так как при выборе формата представления цветовой модели яркостная составляющая включает все блоки, а изменяются только цветные составляющие, соотношение (9) будет иметь следующий вид:

$$D(t) = \sum_{\xi=1}^{m_{mc}} \sum_{\gamma=1}^{n_{mc}} d(t, Y)_{k,l}^{(\xi, \gamma)} + \sum_{\xi=1}^{m_{mc}} \sum_{\gamma=1}^{n_{mc}} \left(\begin{array}{l} \sum_{k=1}^{m_c} \sum_{l=1}^{n_c} d(P, C_r)_{k,l}^{(\xi, \gamma)} + \\ + \sum_{k=1}^{m_c} \sum_{l=1}^{n_c} d(P, C_b)_{k,l}^{(\xi, \gamma)} \end{array} \right). \quad (10)$$

Для формата представления цветного пространства 4:4:4, количество элементов яркостной и хроматических составляющих кадра видеопотока сохранены в полном составе (см. рис. 3), поэтому при $m_c = 2$, $n_c = 2$, объем цифрового описания будет определяться выражением:

$$d(t)_{\text{н\ddot{o}д}}^{(\xi, \gamma)} = \sum_{k=1}^2 \sum_{l=1}^2 \left(d(t, Y)_{k,l}^{(\xi, \gamma)} + d(P, C_r)_{k,l}^{(\xi, \gamma)} + d(P, C_b)_{k,l}^{(\xi, \gamma)} \right).$$

Объем $D(t)$ потока видеоданных в расчете на один кадр соответствует выражению (10). Количество строк и столбцов в макроблоке остается неизменным, т.е. $m_c = 2$, $n_c = 2$ для хроматических составляющих C_r и C_b . Тогда выражение (10) примет вид:

$$D(t) = \sum_{\xi=1}^{m_{mc}} \sum_{\gamma=1}^{n_{mc}} \left(\begin{array}{l} \sum_{k=1}^2 \sum_{l=1}^2 \left(d(t, Y)_{k,l}^{(\xi, \gamma)} + d(P, C_r)_{k,l}^{(\xi, \gamma)} + \right. \\ \left. + d(P, C_b)_{k,l}^{(\xi, \gamma)} \right) \end{array} \right).$$

Для формата представления цветного пространства 4:2:2, горизонтальная выборка хроматических составляющих изображения уменьшится в два раза по сравнению с яркостной составляющей. В этом случае объем цифрового описания структурной единицы определяется как:

$$d(t)_{\text{н\ddot{o}д}}^{(\xi, \gamma)} = d(t, Y)_{k,l}^{(\xi, \gamma)} + \sum_{l=1}^2 \left(d(P, C_r)_{k,l}^{(\xi, \gamma)} + d(P, C_b)_{k,l}^{(\xi, \gamma)} \right).$$

С учетом особенностей формата представления цветного пространства 4:2:2, когда из макроблоков обеих цветных составляющих исключается вторая строка,

т.е. $m_c = 1$, $n_c = 2$, выражение (10) для объема $D(t)$ потока в расчете на кадр примет вид

$$D(t) = \sum_{\xi=1}^{m_{mc}} \sum_{\gamma=1}^{n_{mc}} \left(\begin{array}{l} \sum_{k=1}^2 \sum_{l=1}^2 \left(d(t, Y)_{k,l}^{(\xi, \gamma)} + d(P, C_r)_{1,l}^{(\xi, \gamma)} + \right. \\ \left. + d(P, C_b)_{1,l}^{(\xi, \gamma)} \right) \end{array} \right).$$

При использовании формата цветного пространства 4:2:0 хроматические составляющие C_r и C_b цветовой модели представлены одним блоком первой четверти, в макроблоке отбрасывается вторая строка и второй столбец, т.е. $m_c = n_c = 1$ (см. рис. 3). Для данного формата объем цифрового описания структурной единицы определится следующей формулой:

$$d(t)_{\text{н\ddot{o}д}}^{(\xi, \gamma)} = d(t, Y)_{k,l}^{(\xi, \gamma)} + d(P, C_r)_{1,1}^{(\xi, \gamma)} + d(P, C_b)_{1,1}^{(\xi, \gamma)}$$

Выражение $D(t)$ для формата представления цветного пространства 4:2:0 примет вид

$$D(t) = \sum_{\xi=1}^{m_{mc}} \sum_{\gamma=1}^{n_{mc}} \left(\begin{array}{l} \sum_{k=1}^2 \sum_{l=1}^2 \left(d(t, Y)_{k,l}^{(\xi, \gamma)} + d(P, C_r)_{1,1}^{(\xi, \gamma)} + \right. \\ \left. + d(P, C_b)_{1,1}^{(\xi, \gamma)} \right) \end{array} \right).$$

Таким образом, получены выражения для оценки объема передаваемого видеопотока в расчете на один кадр Р-типа с учетом различных цифровых форматов представления цветного пространства и выбора типа обработки блоков.

Выводы

Разработана методологическая база для управления битовой скоростью сжатого видеопотока. При этом:

- разработана схема обработки блоков яркостной составляющей с помощью двух способов: ДКП как I-тип или ДИКМ как Р-тип. Цветоразностные составляющие будут обрабатываться только с помощью ДИКМ;

- определено количество бит, затраченное на представление одного блока и макроблока для всех составляющих цветовой модели при полиадическом кодировании для блоков Р- и I-типа;

- введено понятие структурной единицы кадра, которая формируется на основе макроблоков по одному от каждой составляющей цветовой модели, и определен ее объем для трех вариантов реализации цифрового формата видео;

- определен общий объем $D(t)$ потока данных в расчете на один кадр в зависимости от выбора цифрового формата макроблоков цветовой модели;

Данные выражения позволяют провести оценку битовой скорости видеопотока для одного кадра с учетом изменения таких параметров как цифровой

формат цветового пространства, размер и тип блоков в кадре, итоговое кодирование данных.

Полученные выражения были сформированы для предсказанных P-кадров, так как большая часть видеоданных кодируется именно в этом формате, что позволяет устранять временную избыточность и значительно уменьшать количество передаваемой информации.

Таким образом, созданная методологическая база позволяет построить систему управления битовой скоростью видеопотока, что особенно актуально в сетях передачи данных с низкой пропускной способностью канала при малом времени обработки видеоизображений или ограниченных вычислительных возможностях устройств обработки, например реализация служб IP-TV или on-line конференций, передача данных по беспроводным Wi-Fi сетям и т.д.

Литература: 1. *Ричардсон Ян*. Видеокодирование. H.264 и MPEG-4 - стандарты нового поколения. М.: Техносфера, 2005. 368 с. 2. *Харченко Н.А.* Метод компрессии видеопотока на основе полиадического кодирования предсказываемых кадров / Н.А. Харченко, В.Н. Кривонос // Радио-

электроника и информатика. 2013. №1. С. 21–28. 3. *Харченко Н.А.* Метод реконструкции предсказываемых кадров в телекоммуникационных технологиях кодирования видеотрафика / Н.А. Харченко // Сучасна спеціальна техніка. К.: МВС ДНДІ, науково-практичний журнал. 2012. Вип. 4. С. 17–25.

Поступила в редколлегию 12.08.2013

Рецензент: д-р техн. наук, проф. Сидорович О.Е.

Баранник Владимир Викторович, д-р техн. наук, профессор, начальник кафедры Харьковского университета Воздушных Сил. Научные интересы: обработка информации в инфокоммуникационных системах. Адрес: Украина, 61166, Харьков, ул. Сумская, 77/79, тел. (057) 704-96-53. E-mail: barannik_v_v@mail.ru.

Харченко Наталия Андреевна, инженер 2 категории ХНУРЭ. Научные интересы: обработка и сжатие видеоданных. Адрес: Украина, Харьков, пр. Ленина 14, тел. (057) 702-14-29.

Бекиров Али Энверович, соискатель Харьковского университета Воздушных Сил. Научные интересы: обработка информации в инфокоммуникационных системах. Адрес: Украина, 61166, Харьков, ул. Сумская, 77/79, тел. (057) 704-96-53.