

УДК 621.39

Р.И. Акимов

ТЕХНОЛОГИЯ КОДИРОВАНИЯ ПАКЕТОВ ПРЕДСКАЗАННЫХ КАДРОВ В ИНФОКОММУНИКАЦИОННЫХ СИСТЕМАХ

В статье рассмотрен вариант повышения степени сжатия видеоинформационного потока в инфокоммуникационных системах в условиях обеспечения заданной целостности информации. Изложены основные компоненты разработки технологии кодирования пакета предсказанных кадров. Предложены аналитические выражения для оценки максимального количества двоичных разрядов на представление пакета предсказанных кадров.

Ключевые слова: кодирование предсказанных кадров, межкадровая апертура.

У статті розглянуто варіант підвищення ступеня стиснення потоку відеоінформації в інфокомунікаційних системах в умовах забезпечення заданої цілісності інформації. Викладено основні компоненти розробки технології кодування пакету передбачених кадрів. Запропоновано аналітичні вирази для оцінки максимальної кількості двійкових розрядів на представлення пакету передбачених кадрів.

Ключові слова: кодування передбачених кадрів, міжкадрова апертура.

The ground of the variant of an increase of a degree of video information stream compression is carried out in the infocommunication systems in the conditions of the information set integrity providing. Basic components of the development of the technologies of encoding of package of the predicted shots are stated. Analytical expressions are built for the estimation of maximal amount of binary digits on the presentation of a package of the predicted shots.

Keywords: encoding of the predicted shots, interframe aperture

Повышение эффективности функционирования железнодорожного транспорта связано с улучшением качества контроля за состоянием и перемещением грузов. Здесь решающую роль выполняют системы объективного контроля. В свою очередь, основная информация для таких систем поступает на основе видеоинформационных датчиков. Такая ситуация характеризуется наличием актуальной проблематики исследований, которая диктуется наличием противоречия. С одной стороны – это требование относительно повышения оперативности и достоверности предоставления видеоинформации, что ведет к увеличению их объемов, а с другой – ограниченные пропускные способности и быстродействие удаленных узлов сбора видеоинформации, что приводит к ограниченной производительности технологической составляющей. Поэтому предлагается развивать технологии кодирования видеоинформационного потока.

Рассмотрим структуру видеопотока. В существующих технологиях компрессии (MPEG-2, MPEG-4) используется кадровый принцип обработки. Вся последовательность кадров разбивается на три типа [1–3]. Кадры первого типа (I-кадры, опорные) сжимаются на основе внутрикадровой обработки. Кадры второго типа (P-кадры, предсказанные) сжимаются с использованием предшествующих I- или P-кадров. Кадры третьего типа (B-кадры) сжимаются с использованием двунаправленного предсказания, т.е. с привлечением предшествующих и после-

дуючих I- и P-кадров. Максимальные искажения происходят на этапе обработки B-кадров. Поэтому наибольший интерес представляют технологии кадровой обработки первых двух типов. Таким образом, одним из вариантов дальнейшего совершенствования технологий компрессии относительно повышения степени компрессии в условиях обеспечения заданной целостности видеоинформации является разработка методов кодирования P-кадров. Значит, *цель исследований* заключается в обосновании и построении технологии кодирования предсказанных кадров.

Разработка технологии кодирования предсказанных кадров

Сжатие P-кадров достигается на основе сокращения двух видов избыточности, а именно: пространственной и временной. Сокращение пространственной избыточности достигается кодированием с преобразованием. Для уменьшения временной избыточности используется кодирование с предсказанием и блоковой компенсацией движения. Эти особенности, а именно наличие подобия в соседних кадрах в видеопоследовательности позволяет дополнительно относительно опорных кадров увеличивать степень сжатия P-кадров. Следовательно, на эффективность всей системы сжатия значимое влияние оказывают предсказанные кадры (P-кадры). Количество P-кадров в видеопоследовательности может быть либо фиксированной частоты, либо определяться автоматически по сложности видеоряда. При этом чем большее количество P-кадров используется в видеопоследовательности, тем выше степень сжатия в целом. Однако существующие технологии кодирования предсказанных кадров обладают рядом существенных недостатков. Первый недостаток заключается в том, что с ростом количества предсказанных кадров резко снижается качество изображения. Это требует использование механизма контроля за ошибками. Ошибка в одном блоке распространится на всю группу последующих P-кадров. Вторым недостатком связан с использованием в современных алгоритмах сжатия трансформированных изображений статистических методов, основанных на кодировании Хаффмана. Здесь существуют такие недостатки, как высокая чувствительность к ошибкам, необходимость в буферном накопителе, невозможность определить степень сжатия до начала кодирования, большие задержки при построении кодового дерева для видеопоследовательностей содержащих большое количество кадров. Поэтому требуется совершенствование технологий обработки предсказанных кадров с заданным уровнем достоверности.

Одно из таких направлений состоит в описании последовательности предсказанных кадров (рис.1) на основе выявления межкадровых апертур.

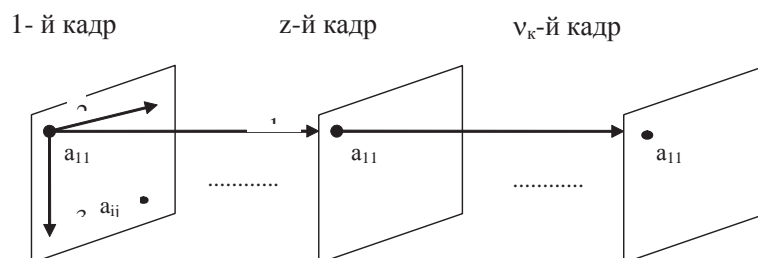


Рис. 1. Последовательность P-кадров

На рис. 1 приняты следующие обозначения: 1 – путь, исключающий межкадровую избыточность; 2, 3 – путь строчной развертки.

Под межкадровой апертурой понимается последовательность со следующими свойствами: элементы, расположенны в пакете P-кадров на одинаковых позициях;

значения элементов находятся в пределах ограниченного динамического диапазона.

Таким образом, апертура содержит не только равные между собой элементы, но и элементы, чьи значения отличаются незначительно. Допуск на диапазон отклонения значений элементов апертуры определяется ее высотой (рис. 2).

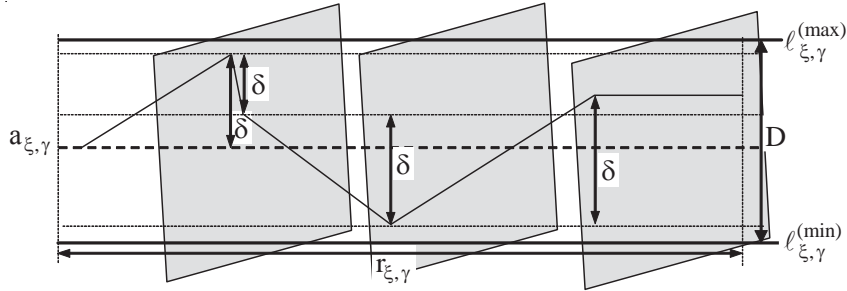


Рис. 2. Схема формирования межкадровой апертуры

Характеристиками межкадровой апертуры являются:

- высота $D_{\xi, \gamma}$ апертуры для $(\xi; \gamma)$ -й последовательности, определяемая как разница между верхним $\ell_{\xi, \gamma}^{(max)}$ и нижним $\ell_{\xi, \gamma}^{(min)}$ значениями динамического диапазона в последовательности элементов предсказанных кадров

$$D_{\xi, \gamma} = \ell_{\xi, \gamma}^{(max)} - \ell_{\xi, \gamma}^{(min)},$$

где $\ell_{\xi, \gamma}^{(min)}$, $\ell_{\xi, \gamma}^{(max)}$ – значения соответственно нижней и верхней границ диаметра $(\xi; \gamma)$ -й апертуры; длина $r_{\xi, \gamma}$ апертуры, вычисляемая как количество подряд расположенных элементов, для которых выполняется условие $a(\tau)_{\xi, \gamma} \in [\ell_{\xi, \gamma}^{(min)}; \ell_{\xi, \gamma}^{(max)}]$, $\tau = 0, r_{\xi, \gamma} - 1$; начальный элемент $a_{\xi, \gamma}$ межкадровой апертуры, относительно которой осуществляется определение величин $\ell_{\xi, \gamma}^{(min)}$ и $\ell_{\xi, \gamma}^{(max)}$, т.е. $\ell_{\xi, \gamma}^{(max)} = a_{\xi, \gamma} + D_{\xi, \gamma} / 2$; $\ell_{\xi, \gamma}^{(min)} = a_{\xi, \gamma} - D_{\xi, \gamma} / 2$.

Обозначим $A^{(\xi, \gamma)}$ как последовательность элементов предсказанных кадров, принадлежащих апертуре, где $A^{(\xi, \gamma)} = \{a(0)_{\xi, \gamma}, a(1)_{\xi, \gamma}, \dots, a(r_{\xi, \gamma})_{\xi, \gamma}\}$, $a(\tau)_{\xi, \gamma} \in [\ell_{\xi, \gamma}^{(min)}; \ell_{\xi, \gamma}^{(max)}]$, $\tau = 0, r_{\xi, \gamma} - 1$.

Обход элементов внутри предсказанного кадра для выявления очередной межкадровой апертуры представлен на рис. 3.

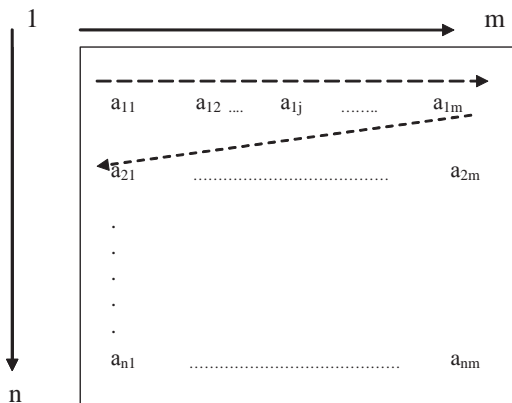


Рис. 3. Схема обхода элементов Р-кадра

---→ -- направление обхода строк для выявления одинаковых элементов.

По окончании обработки элемента a_{mj} для τ -го кадра процесс выявления межкадровых апертур (АМ) продолжается с элемента a_{1j+1} , расположенного на позиции $(1; j+1)$ первого кадра (рис. 3).

В тоже время известные технологии относительно формирования и кодового представления апертур имеют два существенных недостатка, а именно:

1. Формирование апертур, как правило, организуется в режиме заранее заданной высоты, и, следовательно, неравномерной длины. Зависимость неравномерной длины апертуры от степени когерентности фрагментов изображений приводит к: снижению помехоустойчивости; повышению времени на сжатие в случае обработки высоконасыщенных реалистических изображений; уменьшению степени сжатия для изображений с высокой концентрацией резких изменений цвета (яркости), обусловленное сокращением длин апертур и увеличением значений их высот.

2. Обработка апертур базируется на методах с безвозвратной потерей информации. Например, описание апертуры ее длиной, описание апертуры коэффициентами аппроксимирующей функции.

Подходы относительно организации устранения избыточности изображений на основе выявления апертур различаются в зависимости от функции, описывающей элементы, принадлежащие апертуре (*генерирующей апертурной функции (ФАГ)*).

В первом случае обеспечивается обработка с внесением аппроксимирующих искажений в предсказанные кадры, т.е. устраняется психовизуальная и статистическая избыточности. Генерирующая апертурная функция $f'_a(A^{(\xi,\gamma)})$ строится на основе аппроксимации элементов последовательности $A^{(\xi,\gamma)}$ в условиях, когда для высоты апертуры выполняется условие $D_{\xi,\gamma} \geq 1$. Тогда в общем случае происходит несоответствие между исходным и реконструируемым знаниями межкадровой апертуры, т.е. $a'(\tau)_{\xi,\gamma} \neq a(\tau)_{\xi,\gamma}$; $a'(\tau)_{\xi,\gamma} \in \overline{A^{(\xi,\gamma)}}$, $\tau=0, r_\xi-1$, где $A^{(\xi,\gamma)}$, $\overline{A^{(\xi,\gamma)}}$ – последовательности элементов, соответствующие исходной апертуре и апертуре, полученной на основе генерирующей функции $f'_a(A^{(\xi,\gamma)})$:

$$\overline{A^{(\xi,\gamma)}} = f'_a(A^{(\xi,\gamma)}).$$

Здесь $a_{\xi,\gamma}$, $a'_{\xi,\gamma}$ – $(\xi;\gamma)$ -е элементы, принадлежащие соответственно межкадровым последовательностям $A^{(\xi,\gamma)}$ и $\overline{A^{(\xi,\gamma)}}$.

Компактное представление, т.е. выполнение условия $W(P_a) < W(A^{(\xi,\gamma)})$, достигается за счет того, что формируемые параметры P_a генерирующей функции $f'_a(A^{(\xi,\gamma)})$ требуют меньшего количества бит $W(P_a)$, чем битовый объем исходной апертуры $W(A^{(\xi,\gamma)})$.

В случае, когда требуется минимизировать аппроксимирующие искажения, построение генерирующей апертурной функции сопряжено со значительными вычислительными затратами.

Наименьшие временные затраты на реализацию первого подхода достигаются в случае, когда аппроксимирующая функция строится, как среднее значение $a_{\xi,\gamma}$ по элементам межкадровой апертуры $f'_a(A^{(\xi,\gamma)}) = a_{\xi,\gamma}$, а именно

$$\bar{a}_{\xi,\gamma} = \left(\sum_{\tau=0}^{r_{\xi}-1} a(\tau)_{\xi,\gamma} \right) / r_{\xi,\gamma}.$$

Недостаток такого подхода заключается в том, что не обеспечивается контролируемый характер аппроксимирующих искажений. Наоборот минимизация таких искажения требует использования значительных вычислительных ресурсов. Реализация последнего решения ограничивается относительно низким быстродействием аппаратуры дистанционного объективного контроля на ЖД транспорте.

Зависимость коэффициента сжатия k_c от степени h вносимых искажений в процессе выявления апертур в последовательности предсказанных кадров приводится на рис. 4. Оценка проводилась для $\tau=8$. На обработку поступал видеопоток, соответствующий режиму работы стационарной камеры в процессе прохождения железнодорожного состава с грузом.

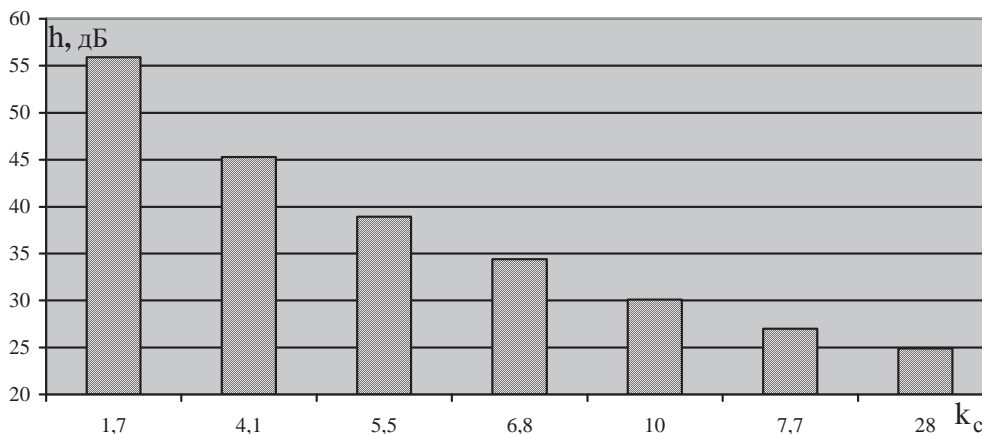


Рис. 4. Зависимость k_c от h для пакета предсказанных кадров

В результате исследования зависимости в виде диаграмм на рис. 4, можно заключить, что уровень сжатия находится в существенной зависимости от степени аппроксимирующих искажений, оцениваемых через пиковое отношение сигнал/шум h . Степень компрессии в режиме получения достоверной информации находится на уровне, не превышающем 4 раз. Следовательно, эффективность обработки предсказанных кадров на основе выявления межкадровых апертур резко снижается для режима обеспечения целостности видеoinформации. В случае наличия в последовательности кадров резких смен сюжетов происходит падение степени сжатия дополнительно на 50–70 %.

Ограниченная возможность относительно обеспечения обработки пакетов Р-кадров заданной длины на основе выявления межкадровых апертур в режиме заданной целостности видеoinформации обусловлена отсутствием механизмов компенсации:

1) повышения информативности, вызванной резким ростом высоты апертуры, соответствующего случаю смены сюжета кадра. Тогда $D_{\xi,\gamma} \rightarrow 2^b$, и соответственно $k_c \rightarrow 1$;

2) наличия дополнительной структурной избыточности, обусловленной отсутствием неоднородных по сюжету кадров изображений. В этом случае длина межкадровой апертуры будет равна длине пакета Р-кадра.

Для второго подхода построение генерирующей апертурной функции предлагается организовывать на базе использования позиционных систем. В этом случае функция $f_a(A^{(\xi,\gamma)})$ создается для следующих условий:

$$D_{\xi,\gamma} \geq 1 \text{ и } x(\tau)_{\xi,\gamma} \neq x(\tau+1)_{\xi,\gamma}, \quad a(\tau)_{\xi,\gamma} \in [\ell_{\xi,\gamma}^{(\min)}; \ell_{\xi,\gamma}^{(\max)}], \quad \tau = \overline{0, r_{\xi,\gamma} - 1}; \quad (1)$$

$$h \geq h_{\text{тр}}. \quad (2)$$

Особенность аппроксимации на основе использования позиционных систем заключается в том, что вся межкадровая апертюра $A^{(\xi,\gamma)}$, состоящая из $r_{\xi,\gamma}$ элементов, заменяется ее кодом-номером $N(r_{\xi,\gamma})$ в допустимом множестве $\Psi(r_{\xi,\gamma})$. Определим объем множества $\Psi(r_{\xi,\gamma})$ как $V(r_{\xi,\gamma})$. Тогда последовательность $A^{(\xi,\gamma)} = \{a(0)_{\xi,\gamma}, a(1)_{\xi,\gamma}, \dots, a(r_{\xi,\gamma})_{\xi,\gamma}\}$ принадлежит множеству $V(r_{\xi,\gamma})$, т.е. $A^{(\xi,\gamma)} \in V(r_{\xi,\gamma})$, если выполняются ограничения $a(\tau)_{\xi,\gamma} \in [\ell_{\xi,\gamma}^{(\min)}; \ell_{\xi,\gamma}^{(\max)}]$ $\tau = \overline{0, r_{\xi,\gamma} - 1}$. Представление межкадровой апертюры на основе предложенного подхода в режиме без потери информации представиться следующим выражением $N(r_{\xi,\gamma}) = f_a(A^{(\xi,\gamma)}; D_{\xi,\gamma}; r_{\xi,\gamma})$, так, что для $\tau = \overline{0, r_{\xi,\gamma} - 1}$:

- 1) $a(\tau)_{\xi,\gamma} \in [\ell_{\xi,\gamma}^{(\min)}; \ell_{\xi,\gamma}^{(\max)}]$;
- 2) $\ell_{\xi,\gamma}^{(\min)} \neq \ell_{\xi,\gamma}^{(\max)}$ и $a'(\tau)_{\xi,\gamma} = a(\tau)_{\xi,\gamma}$.

Здесь $a(\tau)_{\xi,\gamma}$, $a'(\tau)_{\xi,\gamma}$ – (τ) -е элементы, принадлежащие соответственно исходной и восстановленной $(\xi;\gamma)$ -й межкадровой апертюре.

Условие, заданное соотношением (1), означает снятие ограничений на процесс выявления апертюры. Соотношение (2) определяет режим обработки пакета Р-кадров с контролируруемыми аппроксимирующими искажениями.

Следовательно, требуется синтезировать функцию $f_a(A^{(\xi,\gamma)}; D_{\xi,\gamma}; r_{\xi,\gamma})$ для представления межкадровой апертюры $A^{(\xi,\gamma)}$ так, чтобы обеспечить выполнение условия $N(r_{\xi,\gamma}) = f_a(A^{(\xi,\gamma)}; D_{\xi,\gamma}; r_{\xi,\gamma}) \rightarrow \min$ при следующих ограничениях:

$$\begin{cases} r_{\xi,\gamma} = \text{const}; \\ D_{\xi,\gamma} = \text{var} \ \& \ D_{\xi,\gamma} \geq 1; \end{cases} \quad \begin{cases} h \geq h_{\text{тр}}; \\ v_c^{(\phi)}, v_r^{(\phi)} = O(r). \end{cases}$$

Для синтеза генерирующей апертурной функции в заданных условиях заметим, что значения элементов апертюры находятся в ограниченном динамическом диапазоне $\ell_{\xi,\gamma}^{(\min)} \leq a(\tau)_{\xi,\gamma} \leq \ell_{\xi,\gamma}^{(\max)}$. Отсюда вытекает следующее:

1) генерирующая апертурная функция на основе кодообразующего выражения для позиционного числа (ПЧ) примет вид [4]:

$$f_a(A^{(\xi,\gamma)}) = \sum_{\tau=0}^{r_{\xi,\gamma}-1} a(\tau)_{\xi,\gamma} v(\tau)_{\xi,\gamma}, \quad (4)$$

где $v(\tau)_{\xi,\gamma}$ – весовой коэффициент (τ -го элемента ($\xi;\gamma$)-го ПЧ

$$v(\tau)_{\xi,\gamma} = (\ell_{\xi,\gamma}^{(\max)} - \ell_{\xi,\gamma}^{(\min)})^{r_{\xi,\gamma} - (\tau) - 1};$$

2) количество $V^{(\xi,\gamma)}$ различных межкадровых апертур, длиной $r_{\xi,\gamma}$, которое можно составить на основе формирования различных ПЧ равно $V^{(\xi,\gamma)} = (\ell_{\xi,\gamma}^{(\max)} - \ell_{\xi,\gamma}^{(\min)})^{r_{\xi,\gamma}}$;

3) максимальный объем цифрового описания апертуры на основе кодообразующей функции ПЧ оценивается по формуле

$$\log_2 V^{(\xi,\gamma)} = r_{\xi,\gamma} ([\log_2(\ell_{\xi,\gamma}^{(\max)} - \ell_{\xi,\gamma}^{(\min)})] + 1).$$

Откуда суммарный объем W на представление межкадровых апертур, выявленных для всего пакета P -кадров, на основе формирования кода для позиционного числа в режиме переменной апертуры будет равен

$$W = W_{in} + W_{cl}$$

или

$$\begin{aligned} W &= \sum_{\xi=1}^{Q_l} \sum_{\gamma=1}^{Q_c} \log_2 V^{(\xi,\gamma)} + \sum_{\xi=1}^{Q_l} \sum_{\gamma=1}^{Q_c} \log_2 r_{\xi,\gamma} = \\ &= \sum_{\xi=1}^{Q_l} \sum_{\gamma=1}^{Q_c} r_{\xi,\gamma} ([\log_2(\ell_{\xi,\gamma}^{(\max)} - \ell_{\xi,\gamma}^{(\min)})] + 1) + \sum_{\xi=1}^{Q_l} \sum_{\gamma=1}^{Q_c} \log_2 r_{\xi,\gamma}, \end{aligned}$$

где W_{in} , W_{cl} – суммарные объемы цифрового представления соответственно для элементов межкадровых апертур, описываемых на основе выбранной ФАГ, и для описания длин апертур.

В случае, если длина межкадровой апертуры фиксирована и равна, то последнее выражение примет следующий вид:

$$W = \sum_{\xi=1}^{Q_l} \sum_{\gamma=1}^{Q_c} r ([\log_2(\ell_{\xi,\gamma}^{(\max)} - \ell_{\xi,\gamma}^{(\min)})] + 1) + \sum_{\xi=1}^{Q_l} \sum_{\gamma=1}^{Q_c} \log_2 D_{\xi,\gamma}.$$

Здесь $\sum_{\xi=1}^{Q_l} \sum_{\gamma=1}^{Q_c} \log_2 D_{\xi,\gamma}$ – количество разрядов на представление информации

для высот межкадровых апертур.

Таким образом, анализируя полученное соотношение с учетом того, что выполняется неравенство $b \geq ([\log_2(\ell_{\xi,\gamma}^{(\max)} - \ell_{\xi,\gamma}^{(\min)})] + 1)$, где величина b является количеством разрядов, отводимых на представление исходных элементов, можно заключить следующее: суммарный битовый объем пакета P -кадров на основе построения ФАГ, базирующегося на позиционном описании, не будет превышать суммарный цифровой объем до выявления межкадровых апертур.

Выводы

1. Обосновано, что вариантом повышения степени сжатия видеоинформационного потока в инфокоммуникационных системах в условиях обеспечения заданной целостности информации является сокращение временной избыточности на основе выявления межкадровых апертур.

2. Разработана технология кодирования пакета Р-кадров на основе формирования кода-номера позиционным числом, сформированных для межкадровых апертур.

3. Получены аналитические выражения для оценки максимального количества двоичных разрядов на представление пакета Р-кадров, которые показывают, что суммарный битовый объем пакета Р-кадров на основе построения ФАГ, базирующегося на позиционном описании, не будет превышать суммарный цифровой объем до выявления межкадровых апертур.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. *Ватолин Д.* Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. – М. : Диалог-Мифи, 2003. – 381 с.
2. *Ричардсон Я.* Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения / Ян Ричардсон. – Москва : Техносфера, 2005. – 368 с.
3. *Баранник В.В.* Кодирование трансформированных изображений в инфокоммуникационных системах / В.В. Баранник, В.П. Поляков. – Х. : ХУПС, 2010. – 212 с.
4. *Баранник В.В.* Методологические принципы представления апертур во множестве одномерных двухосновных позиционных чисел / В.В. Баранник, Д.С. Кальченко // АСУ и приборы автоматики. – 2011. – Вып. 155. – С. 15–22.

Отримано 24.09.2012