Адаптивное кодирование подвижных изображений на основе трехмерных ортогональных преобразований

Adaptive coding of animated images based on three-dimensional orthogonal conversions

Цветков / Tsvetkov V.
Василий Валерьевич
(3603326@mail.ru)
ЗАО «НПФ «ТИРС»,
начальник отдела.
г. Санкт-Петербург

Устинов / Ustinov A.
Андрей Александрович
ФГКВОУ ВПО «Военная академия связи
имени Маршала Советского Союза
С. М. Буденного» МО РФ,
старший научный сотрудник.
г. Санкт-Петербург

Оков / Okov I.
Игорь Николаевич
ФГКВОУ ВПО «Военная академия связи
имени Маршала Советского Союза
С. М. Буденного» МО РФ,
старший научный сотрудник.
г. Санкт-Петербург

Ключевые слова: алгоритмы сжатия подвижных изображений – animated image shrinking algorithms; трехмерные ортогональные преобразования – three-dimensional orthogonal conversions; видеокодек с адаптацией к изменению статистических свойств кодируемых изображений – video codec with adaptation to variation of statistical properties of coded images.

Показаны недостатки методов кодирования подвижных изображений на основе трехмерных ортогональных преобразований. Предложен подход к построению видеокодека на основе ДКП-3 с адаптацией к изменению статистических свойств кодируемых изображений по оси времени. Осуществлена математическая постановка задачи поиска оптимальной перестановки кодируемых фрагментов подвижного изображения. Приведены результаты имитационного моделирования.

Disadvantages of encoding techniques of animated images based on three-dimensional orthogonal conversions are shown. Approach to video codec creation based on DKP-3 with adaptation to varying statistical properties of coded images along the timeline is offered. Mathematical problem of search of optimum bias of animated image coded fragments is defined. Simulation modeling results are specified.

Введение

Кадры подвижного изображения характеризуются как внутрикадровой или пространственной избыточностью, так и межкадровой или временной избыточностью. Как правило, в известных стандартах сжатия H.263, H.264 для устранения внутрикадровой избыточности используют какое-либо декоррелирующее преобразование, например, двумерное ДКП. Для устранения межкадровой избыточности используют межкадровое предсказание на основе передачи векторов движения. В способах кодирования на основе ДКП-3 [1–7] устранение внутрикадровой и межкадровой избыточности осуществляется путем декорреляции пикселей исходного изображения как по пространственным координатам x и y, так и по оси времени. В результате декорреляции большая часть коэффициентов ДКП-3 оказывается нулевой или близкой к нулю, что и обеспечивает уменьшение требуемого количества бит, необходимых для кодирования коэффициентов ДКП-3. Однако, как показывают практические исследования, декоррелирующее преобразование на основе косинусных функций является оптимальным только для заданного класса изображений. Этот класс изображений ограничивается низкочастотными (с малым числом мелких деталей) изображениями и, в случае подвижного видео, изображениями с медленным изменением сцен при переходе от кадра к кадру. На практике неоптимальность используемых преобразований, особенно во временной области, проявляется в сохранении большого числа ненулевых квантованных коэффициентов преобразования, что, в свою очередь, снижает достигаемый коэффициент сжатия. Устранение этого недостатка возможно на основе реализации процедуры адаптации в процессе кодирования двумя путями. Первый состоит в изменении используемого декоррелирующего преобразования, которое бы учитывало динамику изменения межкадровых различий. Второй состоит в изменении свойств входных данных, над которыми выполняется фиксированное декоррелирующее преобразование. При

этом преобразование входных данных выполняют так, чтобы привести их к виду, оптимальному для используемого декоррелирующего преобразования.

Использование первого способа является довольно проблематичным. Это обусловлено не только необходимостью решения сложной задачи вычисления оптимального преобразования, но и необходимостью передачи больших объемов данных, описывающих полученное преобразование, декодирующему устройству по каналу связи с ограниченной пропускной способностью. В связи с этим в статье рассмотрен подход, основанный на изменении свойств входных данных при сохранении неизменным используемого дискретного косинусного преобразования в качестве декоррелирующего.

Структура кодека на основе ДКП-3 с адаптацией к изменению статистических свойств кодируемых изображений по оси времени

Изменение свойств входных данных предлагается осуществлять на основе изменения порядка следования фрагментов коэффициентов, вычисленных на основе двумерного ДКП по оси времени.

Реализацию данной идеи можно пояснить по схеме, представленной на рис. 1. На вход кодера поступают домены в виде трехмерных массивов пикселей размером $n \times n \times L$. Затем над каждым из L фрагментов домена вычисляют КДКП по координатам x и y, т.е. выполняют операцию ДКП-2. Данную операцию выполняют в блоке 11. Выполнение ДКП-2 осуществляется, как описано в [8].

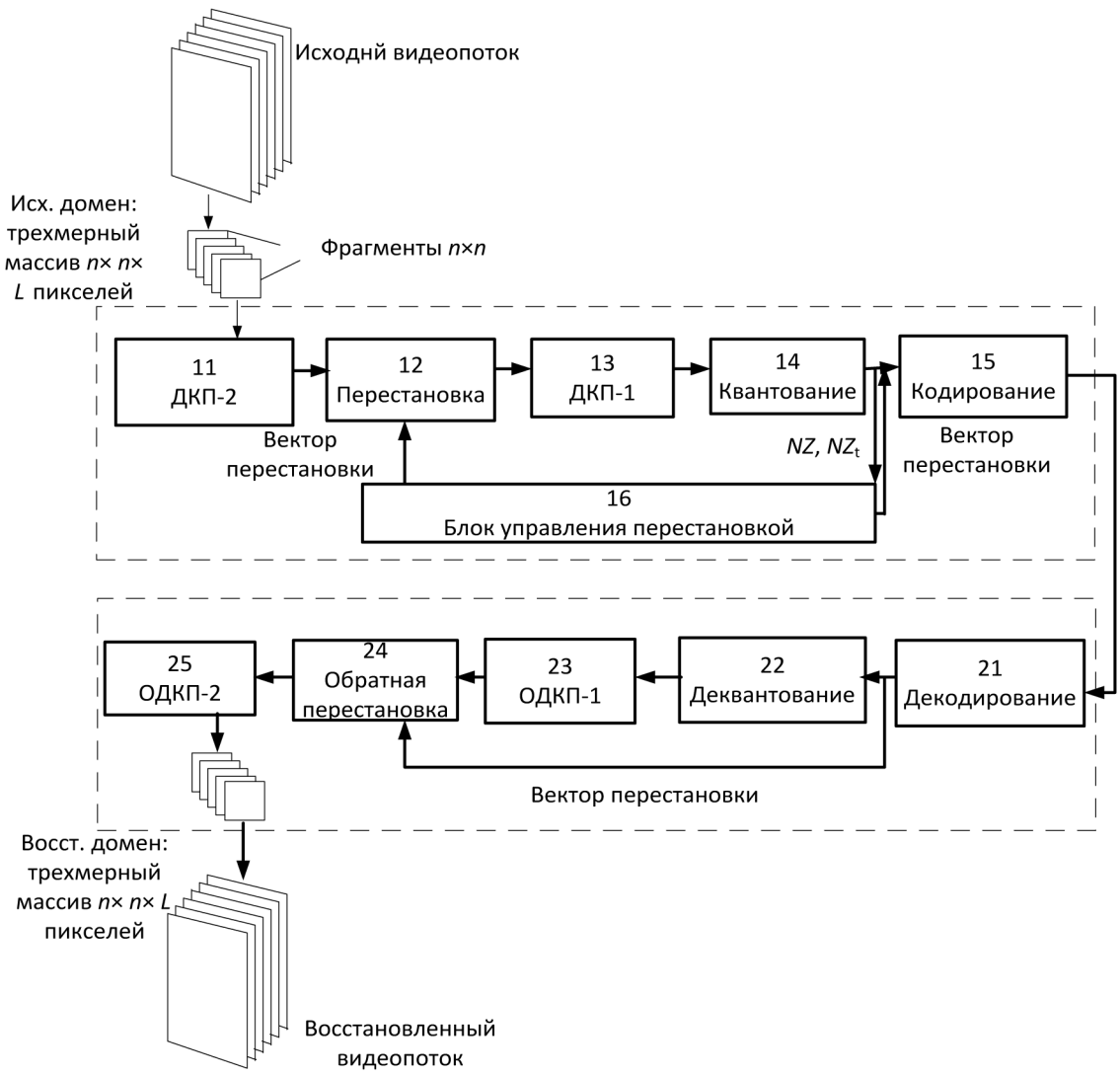

Рис. 1. Структурная схема видеокодека на основе ДКП-3 с оптимальной перестановкой коэффициентов дискретного косинусного преобразования

Далее фрагменты вычисленных коэффициентов ДКП по координатам *x* и *y* поступают на блок перестановки. Перестановка фрагментов выполняется так, чтобы число ненулевых коэффициентов ДКП, полученных после выполнения одномерного ДКП, по оси времени было минимальным. Данную операцию выполняют в блоке 12. Перестановка фрагментов выполняется на основе вектора перестановки, который вычисляется в боке 16 в зависимости от параметров *NZ* и *NZt*, получаемых с выхода блока квантования. После этого выполняют операцию ДКП по времени в блоке 13. Полученные КДКП квантуют в блоке 14 и кодируют в блоке 15.

На приеме декодируют коэффициенты ДКП и вектор перестановки. Данную операцию выполняют в блоке 21. Затем деквантуют коэффициенты ДКП в блоке 22. Над деквантованными коэффициенты ДКП выполняют операцию ОДКП по времени в блоке 23. После этого выполняют обратную перестановку фрагментов коэффициентов ДКП по пространственным координатам x и y в блоке 24. Далее восстанавливают домен размером $n \times n \times L$ пикселей путем вычисления коэффициентов ОДКП по пространственным координатам *x* и *y* в блоке 25 и в результате восстанавливают исходный видеопоток.

Рассмотренные выше операции являются стандартизованными и широко используются во многих видеокодеках. Исключение составляет блок вычисления оптимальной перестановки. Для его реализации рассмотрим математическую постановку задачи поиска оптимальной перестановки.

Математическая постановка задачи поиска оптимальной перестановки кодируемых фрагментов подвижного изображения

В соответствии с описанной выше структурной схемой кодека на основе ДКП-3 на вход кодера последовательно поступают трехмерные массивы пикселей исходного изображения размером $n \times n \times L$. Каждый массив пикселей обрабатывается независимо от других массивов, поэтому достаточно рассмотреть математическую постановку задачи поиска оптимальной перестановки кодируемых фрагментов только для одного массива.

Представим каждый фрагмент массива изображения размером $n \times n \times L$ пикселей множеством из L матриц $\{A_1, A_2, \dots, A_L\}$ размером $n \times n$ каждая и осуществим формальное описание операций ДКП-3 и квантования полученных коэффициентов.

Для удобства описания введем следующие обозначения:

$A_{n \times n \times L}$ – трехмерная матрица, сформированная из L фрагментов размером $n \times n$ каждый. Каждый элемент данной матрицы a(i,j,k), i,j=1,2, ..., n; k=1,2, ..., L – есть значение матрицы $A_{n \times n \times L}$ с координатами (i,j,k);

A(:,:,k) – двумерная матрица k-го слоя трехмерной матрицы $A_{n \times n \times L}$ (см. рис. 2);

A(i,j,:) – вектор строка, сформированная из элементов трехмерной матрицы $A_{n \times n \times L}$ с координатами (i,j,1), (i,j,2), ..., (i,j,L).

В дальнейшем введенные обозначения будем использовать для любой другой трехмерной матрицы.

Выполнение ДКП-3 опишем в виде последовательного выполнения операций двумерного (ДКП-2) и одномерного (ДКП-1) преобразований [6].

Выполнение ДКП-2 над каждым i-ым слоем матрицы $A_{n \times n \times L}$ запишем в виде:

$$B(:,:,i) = D^T A(:,:,i) D, \; i=1,2, \dots, L \qquad (1)$$

где $B(:,:,i)$ – i-ая матрица коэффициентов двумерного дискретного косинусного преобразования; D и D^{T} – прямая и транспонированная (обратная) матрицы ДКП-2, определяемые массивом векторов ,

$\left\{\frac{1}{\sqrt{n}}, \sqrt{\frac{2}{n}} \cos \frac{(2m-1)(k-1)\pi}{2n}\right\}$, m=1, 2,..., n; k=2, 3,..., n. Так, например, матрица ДКП-2 размером 8x8 элементов имеет вид [8]:

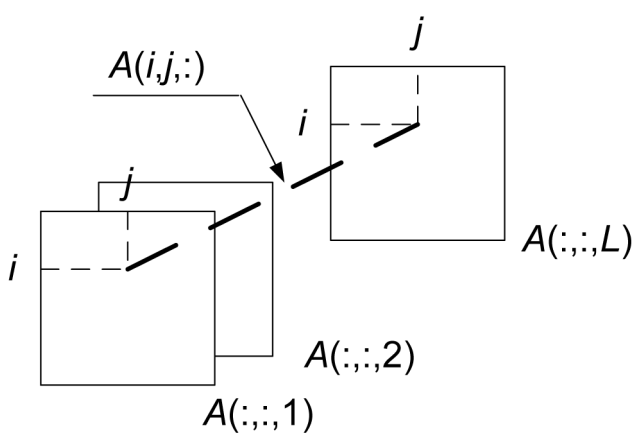

Рис. 2. Элементы трехмерной матрицы

$$D = \begin{bmatrix} 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 \\ 0.4904 & 0.4157 & 0.2778 & 0.0975 & -0.0975 & -0.2778 & -0.4157 & -0.4904 \\ 0.4619 & 0.1913 & -0.1913 & -0.4619 & -0.4619 & -0.1913 & 0.1913 & 0.4619 \\ 0.4157 & -0.0975 & -0.4904 & -0.2778 & 0.2778 & 0.4904 & 0.0975 & -0.4157 \\ 0.3536 & -0.3536 & -0.3536 & 0.3536 & 0.3536 & -0.3536 & -0.3536 & 0.3536 \\ 0.2778 & -0.4904 & 0.0975 & 0.4157 & -0.4157 & -0.0975 & 0.4904 & -0.2778 \\ 0.1913 & -0.4619 & 0.4619 & -0.1913 & -0.1913 & 0.4619 & -0.4619 & 0.1913 \\ 0.0975 & -0.2778 & 0.4157 & -0.4904 & 0.4904 & -0.4157 & 0.2778 & -0.0975 \end{bmatrix}.$$

В результате выполнения (1) трехмерная матрица $A_{n\times n\times L}$ фрагментов исходного изображения преобразуется в трехмерную матрицу $B_{n\times n\times L}$ коэффициентов ДКП-2 такого же размера как показано на рис. 3.

После того как получена трехмерная матрица $B_{n\times n\times L}$ для выполнения ДКП-3 необходимо реализовать операцию ДКП-1 над каждым вектором-строкой $B(i,j,:)$, i=1,2, ..., n; j=1,2, ..., n. Данную операцию формально запишем в виде

$$C(i,j,:) = B(i,j,:)D,\ i=1,2,\ ...,\ n;\ j=1,2,\ ...,\ n, \qquad (2)$$

где $C(i,j,:)$ – вектор-строка коэффициентов ДКП-3.

Итогом выполнения преобразования (2) является трехмерная матрица $C_{n\times n\times L}$ – коэффициентов трехмерного дискретного косинусного преобразования. Наглядно формирование матрицы $C_{n\times n\times L}$ показано на рис. 4.

Перестановку слоев матрицы $B_{n\times n\times L}$ опишем в виде произведения $B(i,j,:)P$, где P – матрица перестановок размером $L\times L$ элементов, элементы которой удовлетворяют ограничениям:

$$P(i,j) \in \{1,0\}; \qquad (3)$$

$$\sum_{j=1}^{L} P(i,j) = 1 \text{ и } \sum_{i=1}^{L} P(i,j) = 1 \qquad (4)$$

i=1,2, ..., L; j=1,2, ..., L.

Тогда трехмерную матрицу коэффициентов ДКП-3 с учетом перестановки слоев матрицы ДКП-2 запишем в виде

$$C_{\text{пер}}(i,j,:) = B(i,j,:)PD,\ i=1,2,\ ...,\ n;\ j=1,2,\ ...,\ n. \qquad (5)$$

В качестве критерия эффективности выбора матрицы перестановок P используем минимум квантованных коэффициентов ДКП-3. В связи с этим для осуществления математической постановки задачи поиска оптимальной перестановки необходимо формализовать вычисление числа квантованных коэффициентов ДКП-3.

Опишем квантование коэффициентов ДКП-3 в виде деления на соответствующие коэффициенты квантования с последующим округлением:

$$W(i,j,k) = round\left(\frac{C(i,j,k)}{M(i,j,k)}\right),\ i,\ j=1,2,\ ...,\ n;\ k=1,2,\ ...,\ L, \qquad (6)$$

где $W(i,j,k)$ – значение квантованного коэффициента трехмерной матрицы $W_{n\times n\times L}$ с координатами $C(i,j,k)$;

$M(i,j,k)$ – значение коэффициента квантования с координатами (i,j,k) трехмерной матрицы квантования – $M_{n\times n\times L}$.

Для вычисления числа квантованных коэффициентов определим матрицу $W'_{n\times n\times L}$, элементы которой удовлетворяют равенству

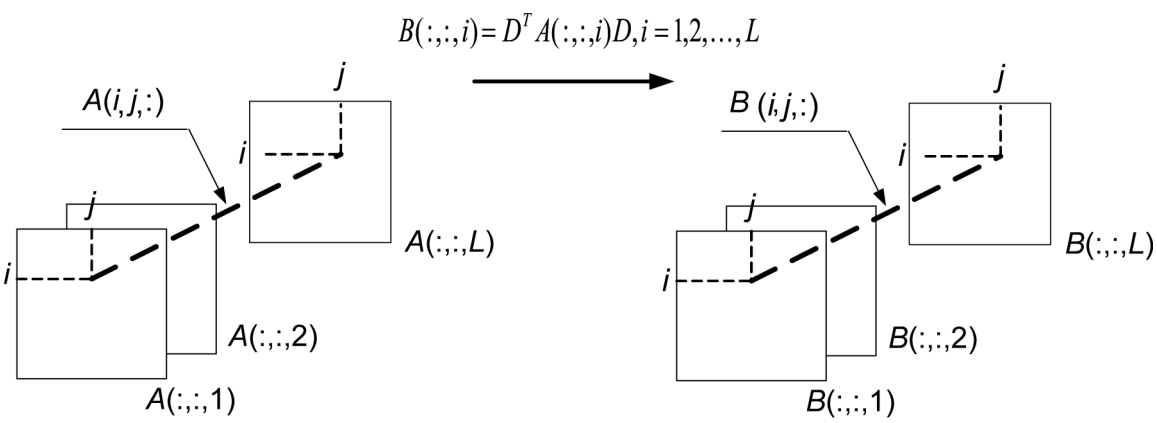

Рис. 3. Иллюстрация выполнения ДКП-2 над трехмерным массивом в матричном виде

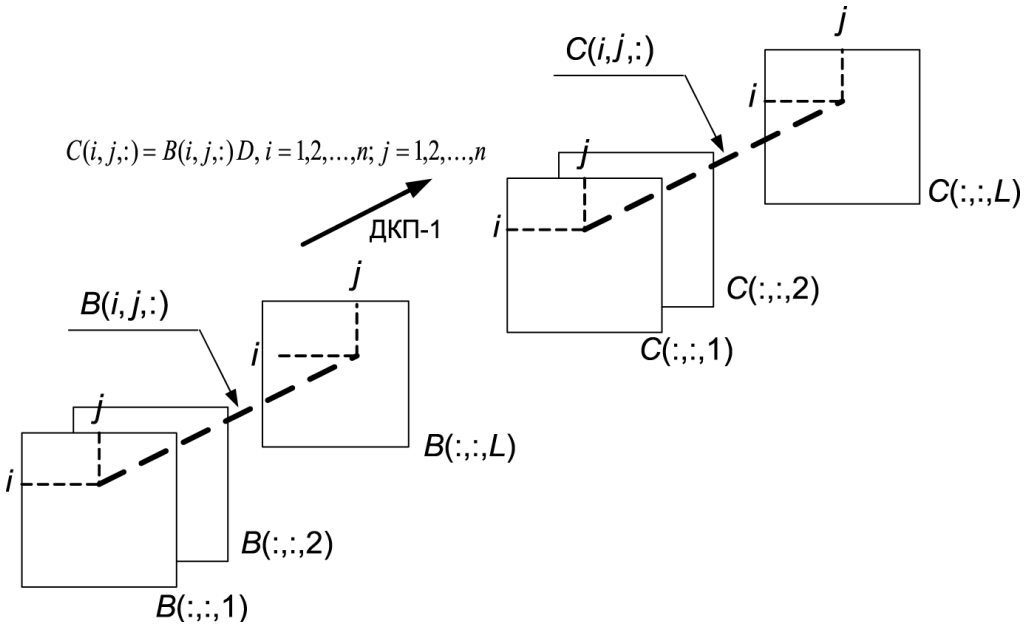

Рис. 4. Иллюстрация выполнения ДКП-1 над трехмерным массивом коэффициентов ДКП-2 (матрицей $B_{n\times n\times L}$)

$$W'(i,j,k) = \begin{cases} 0, \text{если } W(i,j,k) = 0; \\ \frac{1}{W(i,j,k)}, \text{если } W(i,j,k) \neq 0. \end{cases} \quad (7)$$

В этом случае определим матрицу

$$W(i,j,k).*W'(i,j,k)\ ,\ i,j = 1,2,\ ...,\ n;\ k=1,2,\ ...,\ L, \quad (8)$$

где знак (.*) означает перемножение элементов матрицы с одинаковыми индексами.

Матрица (8) характеризуется тем, что она состоит из нулевых и единичных элементов, причем единичные элементы расположены на местах ненулевых квантованных коэффициентов ДКП-3. Тогда для определения их числа необходимо сложить все элементы данной матрицы. В результате сложения получим сумму ненулевых квантованных коэффициентов ДКП-3, т.е. число ненулевых квантованных коэффициентов ДКП-3 есть сумма

$$N_{nz} = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{L} W(i,j,k).*W'(i,j,k). \quad (9)$$

Выражение (9), с учетом ранее введенного обозначения слоя трехмерной матрицы, можно более компактно переписать в виде

$$N_{nz} = \sum_{i=1}^{n}\sum_{j=1}^{n} \vec{1}^T\,(W(i,j,:).*W'(i,j,:))\vec{1}. \quad (10)$$

где $\vec{1}$ – единичный вектор.

В выражении (10) матрицы $W_{n\times n\times L}$ и $W'_{n\times n\times L}$ в соответствии с равенствами (5), (6) и (7) зависят от матрицы перестановок P, т.е. от порядка перестановки кодируемых фрагментов изображения. В связи с этим данные матрицы обозначим как W_P и W'_P Тогда формально постановку задачи поиска оптимальной перестановки кодируемых фрагментов подвижного изображения запишем следующим образом:

$$N_{nz} = arg\max_P \sum_{k=1}^{L} \vec{1}^T(W_P(:,:,k).*W'_P(:,:,k))\vec{1}. \quad (11)$$

Задача (11) состоит в том, чтобы для заданной трехмерной матрицы коэффициентов ДКП-2 – $B_{n\times n\times L}$, трехмерной матрицы коэффициентов квантования $M_{n\times n\times L}$ найти такую матрицу перестановок P, для которой число ненулевых квантованных коэффициентов ДКП-3 будет минимально.

Данная задача была решена в среде Matlab методом случайного поиска с наилучшей пробой. Таким образом, адаптация в процессе кодирования состоит в том, что для каждой новой трехмерной матрицы коэффициентов ДКП-2 ищется новая оптимальная по критерию (11) матрица перестановок.

Результаты имитационного моделирования

Для оценки эффективности предлагаемого способа кодирования и декодирования на основе ДКП-3 было проведено имитационное моделирование на ПЭВМ в среде Matlab.

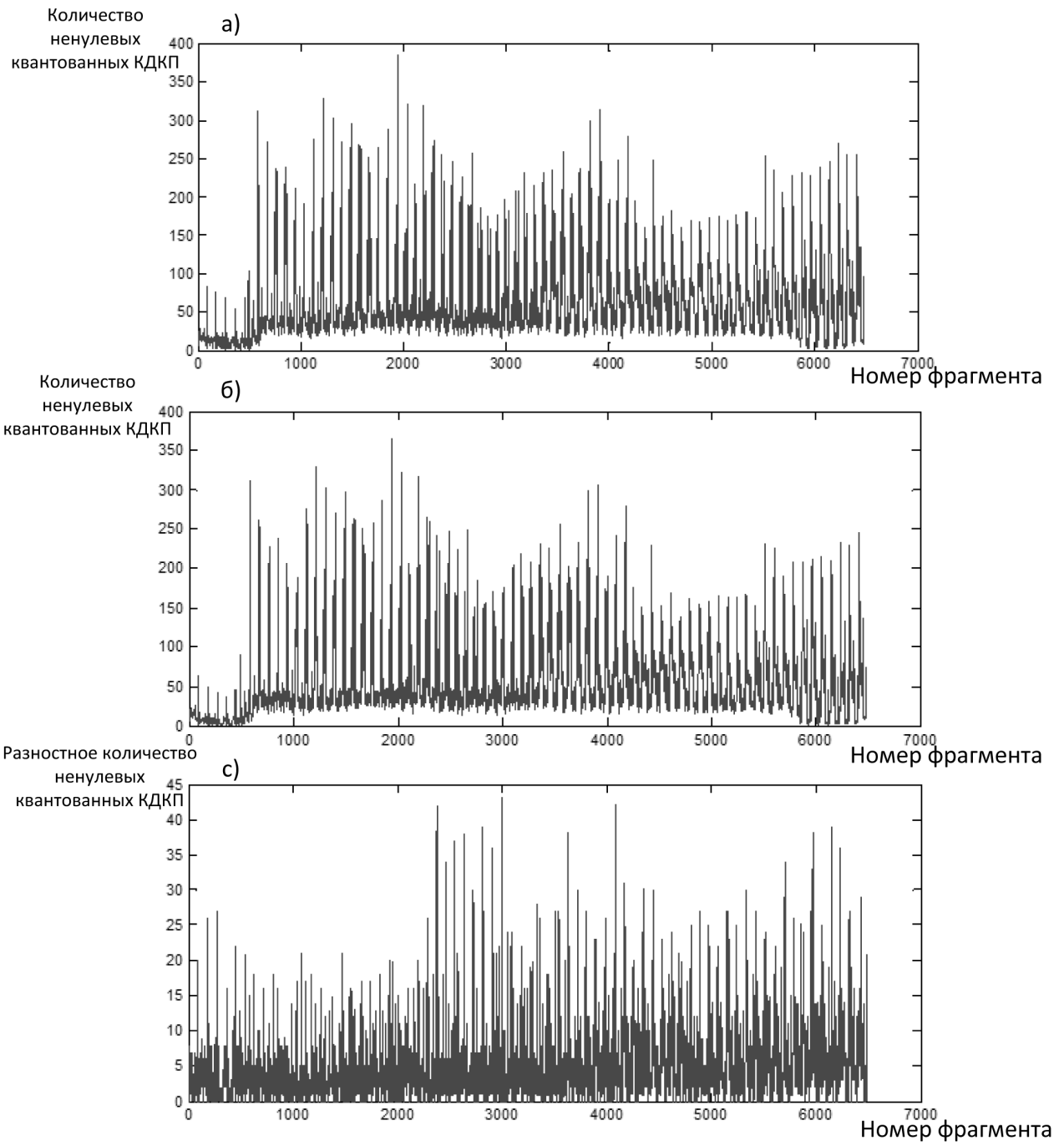

Рис. 4. Зависимости количества ненулевых квантованных коэффициентов ДКП от номера фрагмента:
а) без перестановки фрагментов;
б) с перестановкой фрагментов;
с) величина разности между количеством коэффициентов ДКП, полученных с перестановкой и без перестановки

В качестве показателя эффективности использовался коэффициент снижения количества ненулевых квантованных коэффициентов ДКП при перестановке фрагментов вычисленных коэффициентов ДКП по пространственным координатам x и y относительно количества ненулевых квантованных коэффициентов ДКП, полученных без перестановки.

В качестве исходных подвижных изображений использовался набор тестовых изображений размером 576×720 пикселей в формате *YUV* 4:4:4 и кадровой скоростью 25 кадров/с. Размер исходного домена составил 8×8×16 пикселей.

На рис. 4 показаны характерные зависимости количества ненулевых квантованных коэффициентов ДКП от номера фрагмента: а) без перестановки фрагментов; б) с перестановкой фрагментов; с) величина разности между количеством коэффициентов ДКП, полученных с перестановкой и без перестановки.

Определим коэффициент эффективности как $K_{эф} = \left(\frac{N_{бп} - N_{п}}{N_{бп}}\right) \cdot 100\%$, $N_{бп}$ – суммарное количество ненулевых квантованных коэффициентов ДКП, полученных без перестановки фрагментов; $N_{п}$ – суммарное количество ненулевых квантованных коэффициентов ДКП, полученных с перестановкой фрагментов.

Результаты имитационного моделирования разработанного способа показали, что выигрыш составил 15÷20% по сравнению со случаем без перестановки коэффициентов. Как показали исследования, уменьшение количества ненулевых квантованных коэффициентов ДКП на 15÷20% приводит к такой же величине увеличения коэффициента сжатия при сохранении прежнего качества восстановленных изображений, что подтверждает достижения цели изобретения.

Литература

1. Low-complexity multiple description coding of video based on 3D block transforms / A. Norkin [et al.] // EURASIP Journal on Embedded Systems. – 2007. – № 1. – P. 39.

2. A Versatile Real Time Video Codec Based On Three-Dimensional Discrete Cosine Transform / A. Umbitaliev [et al.] // Ibc 2008, RAI International Congress and Exhibition Centre Amsterdam. The Netherlands. Conference 11–15 September 2008. – P. 386–391.

3. Lee, M. Quantization of 3D-DCT coefficients and scan order for video compression / M. Lee, R. Chan, D. Adjieroh // J. Vis. Commun. Image Represent. – 1997. – № 8. – P. 405–422.

4. Фахми, Ш. С. Адаптивный алгоритм кодирования видеоинформации на основе трехмерного дискретного косинусного преобразования / Ш.С. Фахми, И.А. Зубакин // Изв. вузов России. Радиоэлектроника. – 2010. – Вып. 1. – С. 49–54.

5. Способ кодирования и декодирования видеоинформации на основе трехмерного дискретного косинусного преобразования / Ш.С. Фахми [и др.]. Патент РФ на изобретение № 2375838, заявка № 2008145795 (приоритет от 19.11.2008). – Патентообладатель – Федеральное государственное унитарное предприятие "Научно-исследовательский институт телевидения".

6. Дворников, С. В. Компенсация движения в видеокодеках, использующих трехмерные ортогональные преобразования, на основе оптимальных разбиений кодируемых блоков во временной области / С.В. Дворников, А.А. Устинов, В.В. Цветков // Вопросы радиоэлектроники. Серия техника телевидения. – 2013. – Вып. 2. – С. 98–111.

7. Беляев, Е. А. Сжатие видеоинформации на основе трехмерного дискретного псевдо-косинусного преобразования для энергоэффективных систем видеонаблюдения / Е.А. Беляев, Т.М. Сухов, Н.Н. Шостацкий // Компьютерная оптика. – 2010. – Т. 34. – № 2. – С. 260–272.

8. Ахмед, Н. Ортогональные преобразования при обработке цифровых сигналов [пер. с англ.] / Ахмед Н., Рао К; под ред. И. Б. Фоменко. – М.: Связь, 1980. – 248 с.