

СЖАТИЕ ВИДЕОИНФОРМАЦИИ НА ОСНОВЕ ТРЕХМЕРНОГО ДИСКРЕТНОГО ПСЕВДО-КОСИНУСНОГО ПРЕОБРАЗОВАНИЯ ДЛЯ ЭНЕРГОЭФФЕКТИВНЫХ СИСТЕМ ВИДЕОНАБЛЮДЕНИЯ

Беляев Е.А.¹, Сухов Т.М.², Шостацкий Н.Н.³

¹Санкт-Петербургский институт информатики и автоматизации РАН,

²Балтийский государственный технический университет им. Д.Ф. Устинова,

³ЗАО «Научно-производственная фирма «ТИРС»

Аннотация

В работе рассматриваются методы сжатия видеoinформации на основе трехмерного дискретного косинусного преобразования применительно к системам видеонаблюдения. Для уменьшения вычислительной сложности предложено использовать трехмерное псевдо-косинусное преобразование, которое может быть реализовано без операций умножения, и процедуру квантования, которая может быть реализована без операций деления. Предложен метод временной фильтрации для сжатия видеoinформации в реальном масштабе времени. Приведены результаты сравнения с кодеками на базе стандартов H.264/AVC и MPEG-2.

Ключевые слова: сжатие видеoinформации, трехмерное дискретное косинусное преобразование, временная фильтрация.

Введение

Рост спроса на цифровые устройства видеонаблюдения приводит к необходимости разработки алгоритмов компрессии и декомпрессии видеоданных в реальном масштабе времени. Одним из основных требований к таким устройствам является малое энергопотребление кодера видеoinформации при сохранении эффективности сжатия. Энергопотребление кодера сложным образом зависит от особенностей программной или аппаратной реализации кодера [1]. Однако в общем случае можно считать, что энергопотребление увеличивается с ростом числа операций, выполняемых кодером в единицу времени.

Существующие системы видеонаблюдения чаще всего используют для сжатия видеосигнала кодеки на базе стандартов семейства MPEG (H.264/AVC [2], MPEG-2 [3] и т.д.) и MJPEG [4]. Кодеки на базе стандартов семейства MPEG обеспечивают высокую эффективность сжатия видеоданных, так как при помощи алгоритмов оценки и компенсации движения учитывают межкадровую зависимость отсчетов видеoinформации. Однако данные алгоритмы обладают высокой вычислительной сложностью. Кодеки на базе стандарта MJPEG учитывают только внутрикадровую зависимость отсчетов видеoinформации, и поэтому обладают низкой вычислительной сложностью, но при этом уступают по эффективности сжатия.

Альтернативой перечисленным выше решениям выступают методы сжатия видеоданных на основе трехмерного дискретного косинусного преобразования (ДКП-3), которые обеспечивают компромисс между эффективностью сжатия и вычислительной сложностью. Кроме того, видеоданные, сжатые на основе ДКП-3, более устойчивы к потерям пакетов в каналах связи, в отличие от подходов, использующих алгоритмы оценки движения [5].

Перечисленные особенности делают методы сжатия видеоданных на основе ДКП-3 привлека-

тельными для энергоэффективных цифровых систем видеонаблюдения.

Дальнейшее изложение материала организовано следующим образом. В разделе 1 приведены критерии сравнения кодеков видеoinформации. В разделе 2 рассматривается общая схема кодера на основе ДКП-3. В разделе 3 вводится трехмерное псевдо-косинусное преобразование, которое может быть реализовано без операций умножения, и процедура квантования, которая может быть реализована без операций деления. В разделе 4 вводятся режимы кодирования кубов и методика выбора режимов кодирования для систем, работающих в реальном масштабе времени. Раздел 5 посвящен статистическому кодированию коэффициентов дискретного косинусного преобразования. В разделе 6 приводятся результаты сравнения предложенного алгоритма с кодеками на базе стандартов H.264/AVC и MPEG-2. Раздел 7 посвящен задаче временной предварительной фильтрации для систем, работающих в реальном масштабе времени.

1. Основные критерии сравнения кодеков видеoinформации

Основными характеристиками кодера видеoinформации являются битовая скорость (количество бит, формируемых кодером в единицу времени), уровень искажения восстановленной видеопоследовательности относительно исходной и вычислительная сложность, которую будем измерять количеством операций умножения/деления и сложения/вычитания. Более эффективным сжатием будет обладать кодек, который для заданной битовой скорости обеспечивает меньший уровень искажения.

В качестве меры искажения лучше всего подходят методы субъективной оценки качества, которые основаны на визуальном сравнении изображений. Однако из-за трудоемкости такого сравнения в большинстве работ используются объективные меры, такие как пиковое отношение сигнала к шуму

(PSNR, Peak Signal-to-Noise Ratio) и мера структурной схожести (SSIM, Structural Similarity). Поэтому в данной работе используются именно объективные меры.

Пиковое отношение сигнала к шуму определяется следующим образом:

$$PSNR(d) = 10 \log_{10} \left(\frac{255^2}{d} \right),$$

где d - среднеквадратическая ошибка,

$$d = \frac{1}{K \cdot W \cdot H} \sum_{k \in K} \sum_{(x,y) \in k} (s_k(x,y) - s_k^*(x,y))^2,$$

где H , W – высота и ширина кадра в пикселях, K – число кадров в видеопоследовательности, $s_k(x,y)$, $s_k^*(x,y)$ – значения яркостей пикселей с координатами

x, y в исходном и восстановленном кадре с номером k соответственно.

Мера SSIM учитывает три составляющих искажения: изменение яркости, изменение контраста и изменение структуры. За счет этого, в отличие от пикового отношения сигнала к шуму, эта мера лучше учитывает особенности зрительной системы человека. Более подробно данная мера описана в работе [6].

2. Общая схема кодека на основе ДПКП-3

Общая схема обработки данных в кодере на основе ДПКП-3 представлена на рис. 1. Последовательность действий в кодере/декодере является типовой для данной тематики (см., например, работы [7-11]) и может быть описана следующим образом.

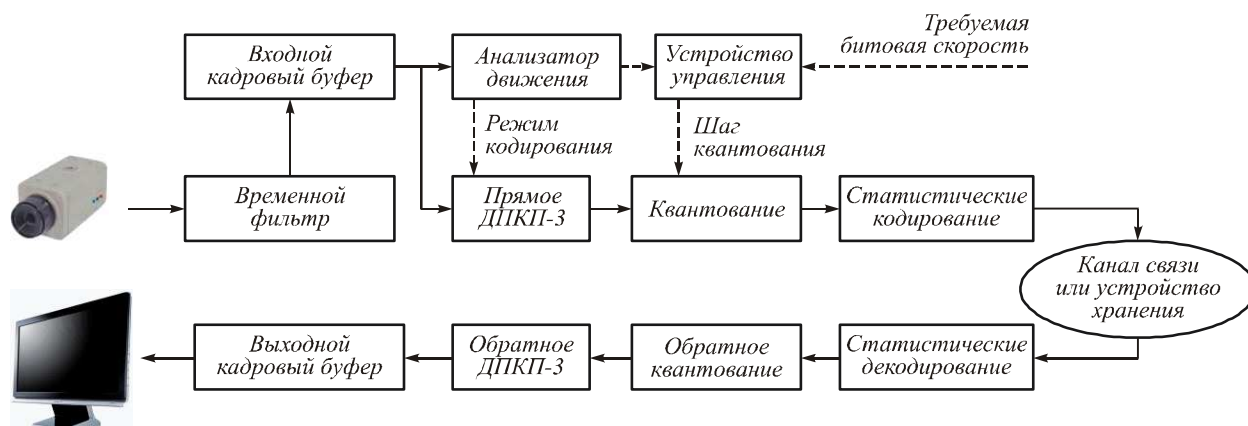


Рис.1. Общая схема кодека на основе ДПКП-3

На стороне кодера *временной фильтр* устраняет шумы, возникающие при формировании кадра в камере видеонаблюдения (см. раздел 7). Необходимое количество фильтрованных кадров накапливается во *входном кадровом буфере* и разбивается на непересекающиеся прямоугольные области, например, на кубы размером $8 \times 8 \times 8$. Следует отметить, что накопление кадров вносит дополнительную задержку кодирования. При использовании ДПКП-3 для кубов размером $8 \times 8 \times 8$ задержка составляет 8 кадров или 320 мс (при кадровой скорости 25 кадров в секунду), что, однако, является приемлемым для систем видеонаблюдения.

Анализатор движения выбирает режим кодирования для каждого куба (см. раздел 4). В зависимости от требуемой битовой скорости *устройство управления* выбирает шаг квантования для каждого куба, после чего выполняется *прямое дискретное псевдо-косинусное преобразование* и процедура *квантования* (см. раздел 3). Квантованные коэффициенты преобразования подаются на вход *статистического кодера* (см. раздел 5), который формирует сжатый поток видеоданных.

На стороне декодера сжатый поток видеоданных обрабатывается *статистическим декодером*, на выходе которого формируются квантованные коэффициенты преобразования. После выполнения *обрат-*

ного квантования и обратного ДПКП-3 восстановленные отсчеты куба помещаются в *выходной кадровый буфер* и передаются на отображающее устройство.

Рассмотрим данную последовательность действий более подробно.

3. Реализация трехмерного преобразования и скалярного квантования

3.1 Трехмерное дискретное косинусное преобразование и квантование с делением

Традиционно прямое дискретное косинусное преобразование для куба размером $N \times N \times N$ задается следующим образом:

$$F(i, j, k) = \sqrt{\frac{8}{N^3}} c(i)c(j)c(k) \sum_{z=0}^{N-1} \sum_{y=0}^{N-1} \sum_{x=0}^{N-1} f(x, y, z) \times \cos \left[\frac{(2x+1)\pi i}{2N} \right] \cos \left[\frac{(2y+1)\pi j}{2N} \right] \cos \left[\frac{(2z+1)\pi k}{2N} \right],$$

где $f(x, y, z)$ – значение яркостной или цветоразностной компонент пикселя с координатами $x, y, z \in [0, \dots, N-1]$, $F(i, j, k)$ – коэффициент преобразования с координатами $i, j, k \in [0, \dots, N-1]$, функция $c(k)$ определяется как

$$c(k) = \begin{cases} \frac{1}{\sqrt{2}}, & k = 0 \\ 1, & k \neq 0. \end{cases}$$

Обратное дискретное косинусное преобразование вычисляется следующим образом:

$$f(x, y, z) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \sqrt{\frac{8}{N^3}} c(i)c(j)c(k) \cdot F(i, j, k) \times \\ \times \cos\left[\frac{(2x+1)\pi i}{2N}\right] \cos\left[\frac{(2y+1)\pi j}{2N}\right] \cos\left[\frac{(2z+1)\pi k}{2N}\right].$$

Как и двумерное ДКП, ДКП-3 обладает следующими свойствами:

- локализация большей части энергии сигнала в небольшом числе коэффициентов преобразования. Данное свойство позволяет исключить из рассмотрения наименее информативные коэффициенты при кодировании с потерями путем квантования [12];

- ДКП-3 представляет собой ортогональное преобразование, которое может быть вычислено путем выполнения одномерного ДКП сначала по строкам, затем по столбцам и по «времени» [13]. Существуют быстрые алгоритмы вычисления ДКП-3, позволяющие выполнить ДКП-3 с использованием порядка $3N^3$ операций умножения и $9N^3 \log_2 N$ операций сложения и вычитания [13-15].

После прямого ДКП-3, для последующего сжатия, следует квантование коэффициентов преобразования. Традиционно используется равномерное скалярное квантование, которое описывается следующей формулой:

$$F_q(i, j, k) = (F(i, j, k) + f \cdot q) / q,$$

где q - шаг квантования, $f \in [0, \dots, 0.5]$.

Восстановление коэффициента преобразования в декодере осуществляется следующим образом:

$$F_r(i, j, k) = F_q(i, j, k) \cdot q.$$

Таким образом, даже с учетом использования быстрого алгоритма вычисления ДКП-3 необходимо выполнить не менее $3N^3$ операций умножения, а также N^3 операций деления при квантовании.

3.2 Трёхмерное дискретное псевдо-косинусное преобразование и квантование без деления

Существенного уменьшения вычислительной сложности можно достичь при помощи перехода от ДКП-3 к трёхмерному псевдо-косинусному преобразованию (ДПКП-3), объединенному с процедурой квантования, реализованной без операции деления. Следует отметить, что аналогичный подход используется в стандарте H.264/AVC для двухмерного преобразования. В настоящей работе предлагается обобщить данный подход на трёхмерный случай.

Рассмотрим сначала процедуру квантования без операции деления. В работе [16] предлагается выполнять квантование отчета x в следующем виде:

$$X_q = (x \cdot A(QP) + f \cdot 2^{20}) / 2^{20}.$$

Обратное квантование вычисляется как:

$$x_r = (X_q \cdot B(QP)) / 2^{20},$$

где $f \in [0, \dots, 0.5]$, $QP = [0, \dots, 31]$ - номер шага квантования, а множители $A(QP)$ и $B(QP)$ задаются с учетом того, что $A(QP) \cdot B(QP) \cong 2^{40}$, как показано в таблице 1.

Таблица 1. Квантование без операции деления

QP	0	1	2	3	4	5
$q(QP)$	2,5	2,8	3,2	3,5	4	4,5
$\frac{A(QP)}{676}$	620	553	492	439	391	348
$\frac{B(QP)}{676}$	3881	4351	4890	5481	6154	6914

Остальные значения коэффициентов $A(QP)$ и $B(QP)$ для $QP > 5$ выбираются из условия, что $2A(QP+6) = A(QP)$ и $B(QP+6) = 2B(QP)$. В связи с этим, в работе [17] предлагается в памяти кодера и декодера хранить только значения коэффициентов $A(QP)$ и $B(QP)$, соответствующих $0 \leq QP \leq 5$. При этом процедура квантования модифицируется следующим образом:

$$X_q = (x \cdot A(q_M) + f \cdot 2^{20+q_E}) / 2^{20+q_E}. \quad (1)$$

Обратное квантование вычисляется как:

$$x_r = (X_q \cdot B(q_M)) / 2^{20+q_E},$$

где $q_M = QP \bmod 6$, $q_E = \lfloor QP / 6 \rfloor$. В результате, квантование может быть реализовано с использованием операции умножения и арифметического сдвига.

Выполненное таким образом квантование соответствует традиционному скалярному квантованию вида:

$$X_r = (x + f \cdot q(QP)) / q(QP),$$

где значение $q(QP)$ связано с номером шага квантования QP , как показано в таблице 1, с учетом того, что $q(QP+1) \approx 1,12 \cdot q(QP)$.

Теперь перейдем к реализации дискретного псевдо-косинусного преобразования без использования операций умножения. Для простоты изложения рассмотрим случай одномерного преобразования над вектором-столбцом длины 8. Обозначим исходный вектор через x , а результат преобразования через X . Тогда дискретное косинусное преобразование можно вычислить следующим образом:

$$X = T \cdot x,$$

где матрица T представляется в следующем виде:

$$T = \begin{pmatrix} g & g & g & g & g & g & g & g \\ a & b & c & d & -d & -c & -b & -a \\ e & f & -f & -e & -e & -f & f & e \\ b & -d & -a & -c & c & a & d & -b \\ g & -g & -g & g & g & -g & -g & g \\ c & -a & d & b & -b & -d & a & -c \\ f & -e & e & -f & -f & e & -e & f \\ d & -c & b & -a & a & -b & c & -d \end{pmatrix},$$

где $a = \frac{1}{2} \cos\left(\frac{\pi}{16}\right)$, $b = \frac{1}{2} \cos\left(\frac{3\pi}{16}\right)$ и т.д.

С учетом ортогональности ДКП, $TT^T = I$, результат обратного преобразования вычисляется следующим образом:

$$x_r = T^T \cdot X.$$

Для реализации преобразования без умножений в работах [18,19] предлагается использовать аппроксимацию матрицы T матрицей

$$H = \frac{1}{8} \times \begin{pmatrix} 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 \\ 12 & 10 & 6 & 3 & -3 & -6 & -10 & -12 \\ 8 & 4 & -4 & -8 & -8 & -4 & 4 & 8 \\ 10 & -3 & -12 & -6 & 6 & 12 & 3 & -10 \\ 8 & -8 & -8 & 8 & 8 & -8 & -8 & 8 \\ 6 & -12 & 3 & 10 & -10 & -3 & 12 & -6 \\ 4 & -8 & 8 & -4 & 4 & 8 & -8 & 4 \\ 3 & -6 & 10 & -12 & 12 & -10 & 6 & -3 \end{pmatrix}.$$

Матрица H , с одной стороны, близка к матрице T , а с другой стороны, вычисление произведений $H \cdot x$ и $H^T \cdot X$ может быть выполнено без операций умножения/деления с использованием 32-х операций сложения/вычитания и 10-ти операций арифметического сдвига (см., например, текст программы в работе [20]).

Для выполнения обратного преобразования необходимо учитывать, что

$$H^T \cdot D \cdot H = I,$$

где D - диагональная матрица. В результате, прямое преобразование может быть вычислено следующим образом:

$$X = D \cdot H \cdot x = (H \cdot x) \otimes d, \quad (2)$$

где $d = D \cdot e$, e - единичный вектор-столбец длины 8, а знак \otimes обозначает поэлементное умножение. Обратное преобразование может быть вычислено как:

$$x_r = H^T \cdot X.$$

Как следует из (2), при вычислении прямого преобразования необходимо выполнить 8 умножений. Избежать этого можно путем переноса операции умножения на вектор d в процедуру квантования.

Для этого в формуле (1) вместо множителя $A(q_M)$ используется множитель:

$$A'(q_M, i) = A(q_M) \cdot d_i,$$

который вычисляется заранее.

Распространяя описанный выше подход, трехмерное псевдо-косинусное преобразование (ДПКП-3) может быть вычислено путем применения описанного выше одномерного преобразования по строкам, столбцам и по «времени». При этом при выполнении операции квантования необходимо использовать множитель $A''(q_M, i, j, k)$, значение которого зависит от координат i, j, k коэффициента преобразования в кубе.

3.3 Сравнение преобразований по вычислительной сложности и эффективности сжатия

В таблице 2 приведено количество операций, необходимое для выполнения рассмотренных выше преобразований совместно с квантованием для куба $8 \times 8 \times 8$. Таблица показывает, что ДПКП-3, совместно с квантованием, не использующим операцию деления, требует существенно меньшего количества операций, чем ДКП-3 с традиционным квантованием.

Таблица 2. Количество операций для куба $8 \times 8 \times 8$

Преобразование	Умножения	Деления	Сложения и вычитания
ДКП-3 (RCF [13]) и квантование с делением	4608	512	11136
ДКП-3 (3-D VR [13]) и квантование с делением	2688	512	11136
ДКП-3 [14] и квантование с делением	2304	512	5568
ДКП-3 [15] и квантование с делением	1920	512	6144
ДПКП-3 и квантование без деления	512	0	6144

На рис. 2-3 показаны результаты сравнения ДКП-3 и ДПКП-3 по эффективности сжатия для различных мер искажения (PSNR и SSIM). Из рис. 2-3 следует, что для фиксированной битовой скорости визуальное качество для ДПКП-3 значительно уступает ДКП-3. С учетом того, что ДПКП-3 обладает существенно меньшей вычислительной сложностью, оно является более предпочтительным для энергоэффективных систем видеонаблюдения.

4. Режимы обработки кубов

4.1 Классификация типов кубов

Для упрощения в настоящей работе рассматривается случай, когда видеоданные разбиваются на непересекающиеся кубы $8 \times 8 \times 8$, каждый из которых может быть отнесен к следующим типам:

1. Статический куб.
2. Куб с умеренным движением.
3. Динамический куб.

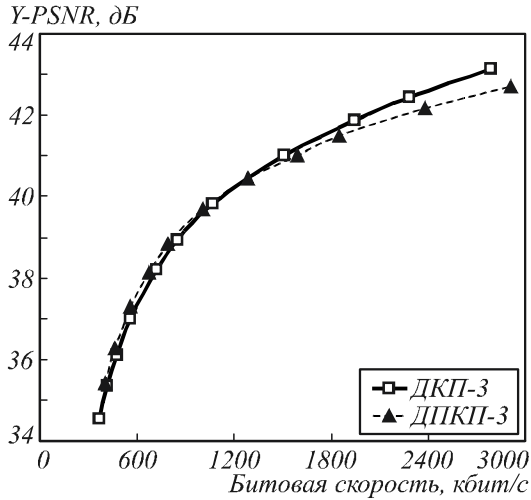


Рис. 2. Сравнение преобразований по пиковому отношению сигнала к шуму

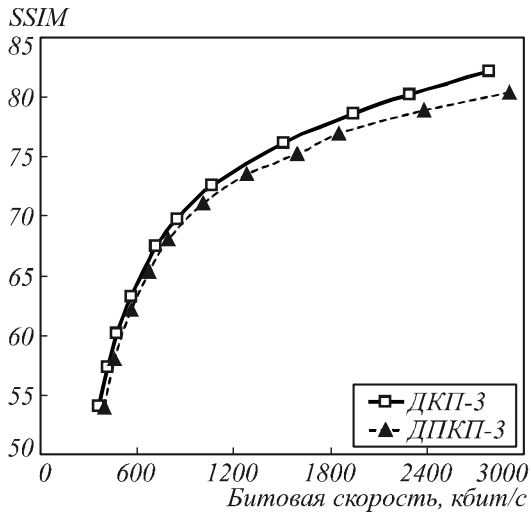


Рис. 3. Сравнение преобразований по мере SSIM

К *статическим кубам* относятся кубы, которые соответствуют неподвижным участкам видеопоследовательности. Для таких кубов дискретное косинусное преобразование и квантование не выполняется, статистическое кодирование не используется. Вместо этого кодер передает один бит, который указывает, что данный куб статический, а декодер не меняет соответствующий участок изображения.

Куб с умеренным движением обрабатывается в соответствии с подходом, описанным в разделе 3. Сначала выполняется ДПКП-3 и квантование, затем выполняется статистическое кодирование.

Если плоскости куба существенно отличаются друг от друга, то выполнение ДПКП-3 становится неэффективным, так как ведет к появлению специфических искажений (см. рис. 4). В этом случае куб классифицируется как *динамический* и вместо ДПКП-3 выполняется двухмерное дискретное псевдо-косинусное преобразование для каждой плоскости куба. Такой подход позволяет устранить возникающие при ДПКП-3 искажения (см. рис. 5).

4.2 Выбор режима обработки куба

Задача выбора режимов обработки кубов для случая, если в качестве меры искажения используется мера PSNR, может быть сформулирована в виде следующей оптимизационной задачи. Для каждого куба с номером i необходимо выбрать такой режим кодирования $m_i \in M$ и шаг квантования $q_i \in Q$, чтобы

$$\begin{aligned} & \text{минимизировать } \sum_i d(m_i, q_i) \\ & \text{при условии, что } \sum_i r(m_i, q_i) \leq R_{\max}, \end{aligned} \quad (3)$$

где $d(m_i, q_i)$ и $r(m_i, q_i)$ - среднеквадратическая ошибка и битовые затраты для куба i при выборе режима обработки m_i и шага квантования q_i , R_{\max} - требуемые битовые затраты на группу кубов.

Решение задачи (3) может быть получено, например, при помощи метода Лагранжевых релаксаций [21]. Однако данный подход обладает высокой вычислительной сложностью. Для систем, работающих в реальном масштабе времени, в настоящей работе предлагается эвристический подход, при котором режим обработки куба выбирается при помощи так называемого *анализатора движения*.

Обозначим через $c(x, y, z)$ и $c'(x, y, z)$ значение яркостной или цветоразностной компоненты кодируемого и предыдущего кубов соответственно, $x, y, z \in [0, \dots, N-1]$. Выбор режима обработки куба основывается на использовании метрик M_1 и M_2 , которые вычисляются следующим образом:

$$M_1 = \max \left(m_1(0, 0), m_1\left(0, \frac{N}{2}\right), m_1\left(\frac{N}{2}, 0\right), m_1\left(\frac{N}{2}, \frac{N}{2}\right) \right),$$

$$M_2 = \max \left(m_2(0, 0), m_2\left(0, \frac{N}{2}\right), m_2\left(\frac{N}{2}, 0\right), m_2\left(\frac{N}{2}, \frac{N}{2}\right) \right),$$

где

$$m_1(a, b) = \max_{z \in \{0, \dots, N-1\}} \sum_{x=a}^{a+\frac{N}{2}-1} \sum_{y=b}^{b+\frac{N}{2}-1} |c'(x, y, z) - c(x, y, z)|,$$

$$m_2(a, b) = \max_{z \in \{1, \dots, N-1\}} \sum_{x=a}^{a+\frac{N}{2}-1} \sum_{y=b}^{b+\frac{N}{2}-1} |c(x, y, 0) - c(x, y, z)|.$$

Если $M_1 < T_1$ и $M_2 < T_1$, то куб классифицируется как статический. Если $M_2 > T_2$, то куб классифици-

руется как динамический. В остальных случаях куб классифицируется как куб с умеренным движением.

Для обеспечения возможности воспроизведения видеоданных с произвольного места, а также для повышения устойчивости передачи по ненадежным каналам связи анализатор движения не классифицирует кубы на одной и той же позиции как статические более чем K раз подряд.



Рис. 4. Сжатие с использованием только трехмерного преобразования



Рис. 5. Сжатие с использованием двумерного преобразования для динамических участков видеоданных

Следует отметить, что особенность метрик M_1 и M_2 заключается в том, что они учитывают наличие движения отдельно в каждой четверти куба, что повышает точность классификации для кубов, находящихся на границах подвижных объектов и статических областей.

5. Статистическое кодирование коэффициентов преобразования

Для формирования битового потока сжатых видеоданных квантованные коэффициенты дискретного псевдо-косинусного преобразования подаются на вход статистического кодера. В настоящей работе применяется подход, приведенный в стандарте

MPEG-2, в котором используется кодирование длин серий путем обхода коэффициентов по «зиг-загу» с последующим кодированием длин серий и амплитуд коэффициентов преобразования при помощи кода Хаффмана. Единственным отличием является порядок обхода коэффициентов для трехмерного случая.

В работе [22] предлагается использовать двумерный порядок обхода для каждой плоскости куба, как это реализовано в стандарте MPEG-2. В работе [23] предлагается использовать более сложный параболический пространственный обход. Однако с точки зрения степени сжатия более эффективно использовать обход, соответствующий убыванию частот появления ненулевых коэффициентов преобразования в кубе (см. рис. 6), так как в этом случае обеспечиваются минимальные битовые затраты при кодировании длин серий.

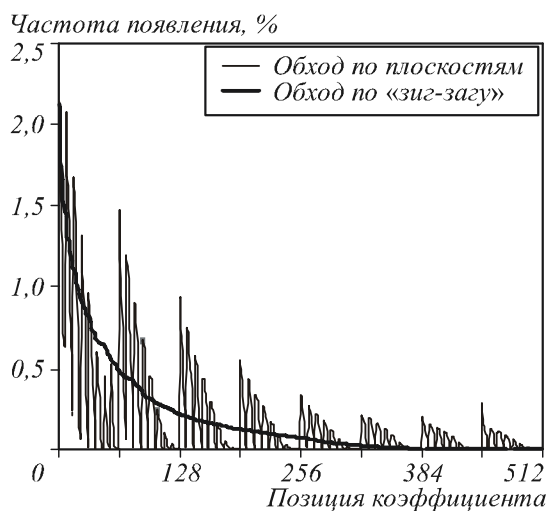


Рис. 6. Частоты появления коэффициентов ДКП

6. Результаты практического сравнения

Практические результаты были получены для известных тестовых видеопоследовательностей [24]: «akiyo», «container», «hall», «mother-daughter», «news» и «silent» длительностью 300 кадров, разрешением 352x288, с кадровой скоростью 30 кадров в секунду (см. рис. 7-12). Данные видеопоследовательности содержат движущиеся объекты, снятые неподвижно установленной камерой, то есть являются типичными для систем видеонаблюдения.

Для сжатия видеопоследовательностей предложенным алгоритмом на основе ДПКП-3 использовались значения порогов анализатора движения $T_1 = 4$, $T_2 = 14$ и период $K = 5$.

Для сравнения перечисленные видеопоследовательности также были сжаты при помощи открытых реализаций стандартов H.264/AVC (JM кодек [25], версии 16.2) и MPEG-2 [26]. В случае H.264/AVC использовался базовый профиль (Base Profile), который обладает наименьшей вычислительной сложностью по сравнению с остальными профилями, которые используют оценку движения.

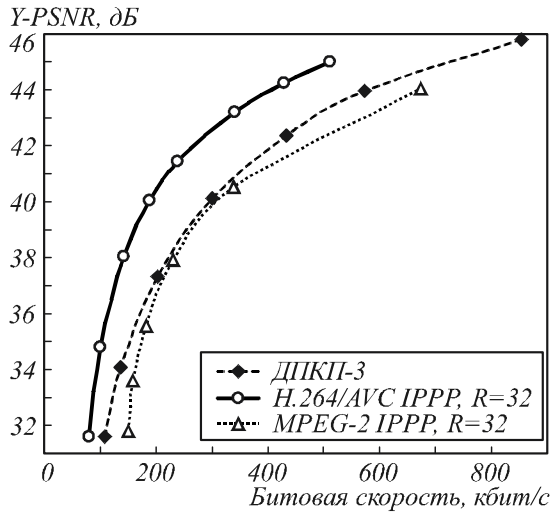


Рис. 7. Результаты сравнения для «akiyo»

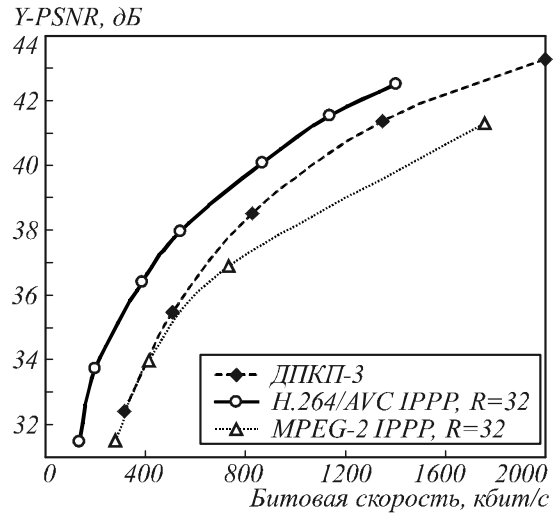


Рис. 8. Результаты сравнения для «container»

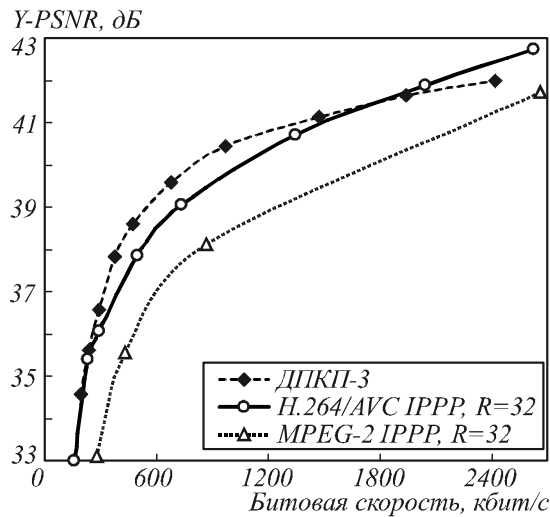


Рис. 9. Результаты сравнения для «hall»

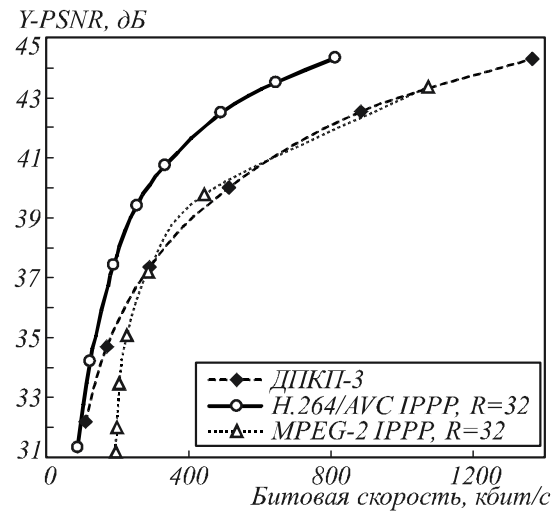


Рис.10. Результаты сравнения для «mother-daughter»

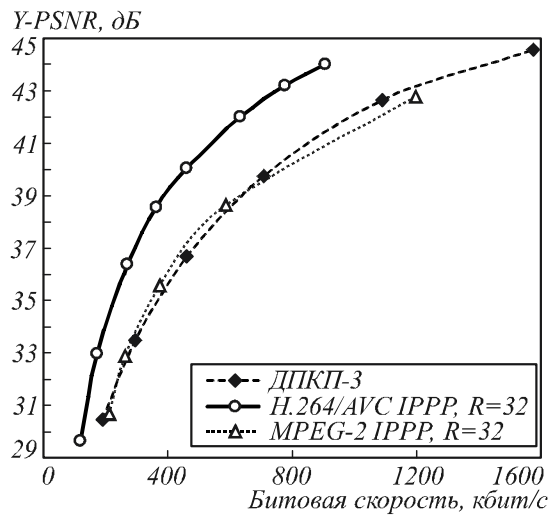


Рис.11. Результаты сравнения для «news»

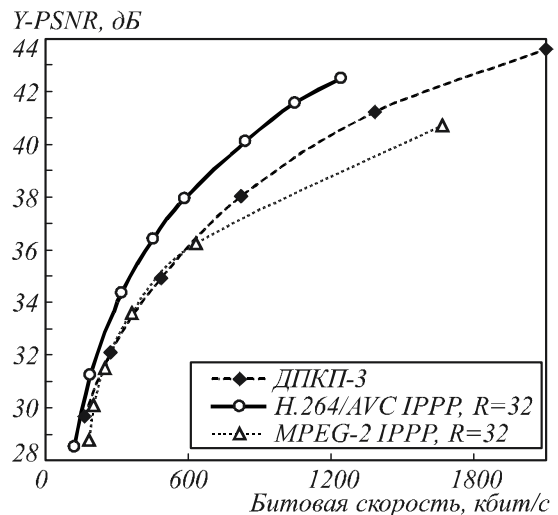


Рис.12. Результаты сравнения для «silent»

В случае MPEG-2 использовался режим кодирования без кадров с двунаправленным предсказанием (B-frames). В обоих случаях использовалась оценка и компенсация движения в радиусе $R = 32$ пикселя.

Приведенные результаты показывают, что для последовательностей, типичных для систем видеонаблюдения, предложенный алгоритм выигрывает у кодека на базе стандарта MPEG-2 и несколько усту-

пает базовому профилю стандарта H.264/AVC. Прогресс относительно H.264/AVC обусловлен использованием в стандарте более совершенного статистического кодера (CAVLC, Context adaptive variable-length coder), который эффективнее учитывает статистические зависимости между квантованными коэффициентами ДКП. В связи с этим можно ожидать, что введение аналогичного статистического кодера в кодек на базе ДПКП-3 позволит значительно улучшить эффективность кодирования.

С учетом того, что рассматриваемый в работе алгоритм не использует оценку движения, не использует операции умножения при выполнении ДКП и операций деления при квантовании, его вычислительная сложность значительно меньше, чем у рассмотренных кодеков на базе стандартов MPEG-2 и H.264/AVC. Поэтому он является более предпочтительным для энергоэффективных систем видеонаблюдения.

7. Временная предварительная фильтрация

В цифровых системах видеонаблюдения видеосигналы, как правило, искажаются шумами, которые могут иметь различную природу возникновения. Эти шумы не только ухудшают качество визуального восприятия видеоданных, но и значительно понижают эффективность сжатия источника видеoinформации. Поэтому подавляющее число современных систем видеонаблюдения перед сжатием использует предварительную фильтрацию видеоданных. В случае ДПКП-3 уменьшение уровня шума также повышает эффективность описанного в подразделе 4.2 эвристического алгоритма выбора режима кодирования куба.

Один из способов устранения шумовой составляющей может быть основан на базе временной предварительной фильтрации. При этом для систем, работающих в реальном масштабе времени, из всего множества методов фильтрации могут быть использованы только те методы, которые сочетают приемлемое качество фильтрации с низкой вычислительной сложностью.

В данной работе предлагается решить данную задачу при помощи рекурсивного фильтра, схема которого представлена на рис. 13. Во-первых, для реализации данного фильтра требуется объем памяти, необходимый для хранения только одного кадра видеопоследовательности. Во-вторых, коэффициенты фильтра $a, b \in [0, \dots, 1]$ могут быть заданы таким образом, чтобы используемые при фильтрации операции умножения заменялись операциями сложения и арифметического сдвига.

С уменьшением значений коэффициентов фильтра a и b понижается шумовая составляющая видеоданных и, следовательно, увеличивается степень сжатия, достигаемая при использовании ДПКП-3.

Однако малые значения коэффициентов приводят к эффекту «размазывания» изображений дви-

жущихся объектов, что ухудшает визуальное восприятие видеoinформации.

Для устранения данного эффекта предлагается использовать анализатор движения, который в зависимости от степени подвижности объектов использует различные коэффициенты фильтра.

В простейшем случае кадр делится на непересекающиеся блоки. Для каждого блока вычисляется средняя абсолютная разность на пиксель. Если данная разность меньше порога T , то блок классифицируется как статический и для пикселей внутри блока используется рекурсивный фильтр с коэффициентами a_l и b_l . В противном случае блок классифицируется как динамический и для пикселей внутри блока используется рекурсивный фильтр с коэффициентами a_h и b_h .

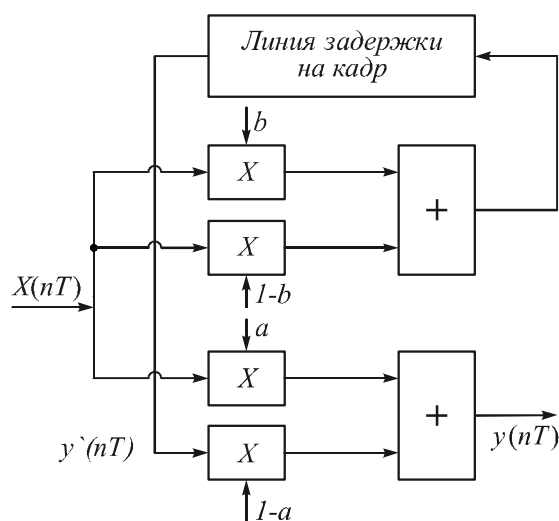


Рис. 13. Схема временной предварительной фильтрации на базе рекурсивного фильтра

Для примера, на рис. 14-16 приведен один кадр видеопоследовательности, сжатой при помощи ДПКП-3 с одинаковым шагом квантования. На рис. 14 приведен кадр без использования предварительной фильтрации. На рис. 15 приведен кадр с использованием предварительной фильтрации с коэффициентами $a = 1/2$, $b = 3/4$. На данном рисунке хорошо заметен эффект «размазывания» изображения идущего в центре кадра человека. На рис. 16 приведен кадр с использованием предварительной фильтрации и анализатора движения с порогом $T = 8$, и коэффициентами фильтров $a_l = 1/2$, $b_l = 3/4$ и $a_h = 7/8$, $b_h = 15/16$. Как видно из рис. 16, использование анализатора движения привело к устранению эффекта «размазывания».

Таким образом, использование предложенной временной предварительной фильтрации позволяет существенно уменьшить битовую скорость при сохранении визуального качества передаваемой видеoinформации (см. значения битовых скоростей на рис. 14-16). При этом предложенный фильтр может быть реализован без использования

операций умножения/деления, что позволяет использовать его в системах, работающих в реальном масштабе времени.



Рис.14. Сжатие на основе ДПКП-3 без использования временной предварительной фильтрации (1746 кбит/с)



Рис.15. Сжатие на основе ДПКП-3 с использованием предложенной временной предварительной фильтрации без анализатора движения (791 кбит/с)



Рис.16. Сжатие на основе ДПКП-3 с использованием предложенной временной предварительной фильтрации с анализатором движения (919 кбит/с)

Выводы

В работе были рассмотрены методы сжатия видеoinформации на основе трехмерного дискретного косинусного преобразования применительно к системам цифрового видеонаблюдения. Предложено использовать трехмерное дискретное псевдо-косинусное преобразование, которое может быть реализовано без операций умножения, и процедуру квантования, которая может быть реализована без операций деления. Предложен эвристический алгоритм, позволяющий выполнять выбор режимов кодирования кубов, а также метод временной предварительной фильтрации для систем, работающих в реальном масштабе времени. Приведены результаты практического использования алгоритма по сравнению с кодеками на базе стандартов H.264/AVC и MPEG-2.

Полученные результаты показывают, что на типичных для систем видеонаблюдения последовательностях предложенный алгоритм на основе ДПКП-3 не уступает по эффективности сжатия алгоритмам, использующим оценку движения. При этом он обладает существенно меньшей вычислительной сложностью, поэтому является более предпочтительным для энергоэффективных систем видеонаблюдения.

За рамками настоящей работы остались вопросы, связанные с функционированием устройства управления, результатами сравнения предложенного кодера для случая передачи видеоданных по каналам связи с потерями пакетов, пространственными пред- и постфильтрацией, а также иерархическим способом кодирования на базе ДПКП-3. Мы планируем рассмотреть эти вопросы в последующих работах.

Благодарности

Работа выполнена при поддержке РФФИ по проектам № 10-08-01071-а и №08-08-00403-а, а также в рамках программы фундаментальных исследований ОНИТ РАН по проекту 2.3.

Литература

1. **He, Zh.** Power-rate-distortion analysis for wireless video communication under energy constraints / Zhihai He, Y. Liang, L. Chen, I. Ahmad and Dapeng Wu // IEEE Transactions on Circuits and Systems for Video Technology. – 2005. – Vol. 15. – P. 645-658.
2. Advanced video coding for generic audiovisual services – ITU-T Recommendation H.264 and ISO/IEC 14496-10 (AVC), 2009.
3. ISO/IEC 13818 (MPEG-2): Generic coding of moving pictures and associated audio information, 1994.
4. ITU-T and ISO/IEC JTC1, «Digital Compression and coding of continuous-tone still images», ISO/IEC 10918-1 – ITU-T Recommendation T.81 (JPEG), 1992.
5. **Ji, X.** Concealment of Whole-Picture Loss in Hierarchical B-Picture Scalable Video Coding / Xiangyang Ji, Debin Zhao and Wen Gao // IEEE Transactions on Multimedia. – 2008. – Vol. 11. – P. 11-22.
6. **Wang, Z.** Image quality assessment: From error visibility to structural similarity / Z. Wang, Alan C. Bovik, Hamid R. Sheikh and Eero P. Simoncelli // IEEE Transactions on Image Processing. – 2004. – Vol. 13, N. 4. – P. 600-612.
7. **Koivusaari, J.J.** Simplified three-dimensional discrete cosine transform based video codec / J.J. Koivusaari and

- J.H. Takala // SPIE Proceedings in Multimedia on Mobile Devices. – 2005. – Vol. 5684. – P. 11–21.
8. **Chan, Y.-L.** Variable Temporal-Length 3-D Discrete Cosine Transform Coding / Yui-Lam Chan and Wan-Chi Siu // IEEE Transactions on Image Processing. – 1997. – Vol. 6, No. 5.
 9. **Servais, M.** Video Compression using the Three Dimensional Discrete Cosine Transform / M. Servais and Gerhard De Jager // Proceedings of the 1997 South African Symposium on Communications and Signal Processing, 1997.
 10. **Chan, R.K.W.** 3D-DCT Quantization as a Compression Technique for Video Sequences / R.K.W. Chan and M.C. Lee // Proceedings of the 1997 International Conference on Virtual Systems and MultiMedia, 1997.
 11. **Furht, B.** An Adaptive Three-Dimensional DCT Compression Based on Motion Analysis / B. Furht, Ken Gustafson, Hesong Huang and Oge Marques // Proceedings of the 2003 ACM symposium on Applied computing, 2003.
 12. **Красильников, Н.Н.** Цифровая обработка изображений – М.: Вузовская книга, 2001. – 320 с.
 13. **Mekky, T.** On the computation of the 3-D DCT / T. Mekky, S. Boussakta and M. Darnell – IEEE International Conference on Electronics, Circuits and Systems, Vol. 3, 2003. – P. 1141 – 1143.
 14. **Chan, S.C.** Fast algorithms for computing the discrete cosine transform / S.C. Chan and K.L. Ho // IEEE Transactions on Circuits Systems II: Analog Digital Signal Process // 1992. – Vol. 39. – P. 185-190.
 15. **Chichyeva, M.A.** Image Block Coding Based on New Algorithms of Shortlength DCT with Minimal Multiplicative Complexity / M.A. Chichyeva, and V.M. Chernov // F. Solina, A. Leonardis (Eds).
 16. **Wiegand, T.** H.26L Test Model Long-Term Number 9 (TML-9) draft0 // ITU-T Q.6/SG16 Document TML9.doc (VCEG-N83d1), 2001.
 17. **Hallapuro, A.** Low complexity transform and quantization – Part 1: Basic Implementation / A. Hallapuro and M. Karczewicz – JVT document JVT-B038, 2002.
 18. **Malvar, H.S.** Low-complexity transform and quantization in H.264/AVC / H.S. Malvar, A. Hallapuro, M. Karczewicz and L. Kerofsky // IEEE Transactions on Circuits and Systems for Video Technology. – 2003. – Vol. 13(7). – P. 598-603.
 19. **Gordon, S.** Simplified Use of 8x8 Transform – Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, doc. JVT-I022, 2003.
 20. **Lin, Y.K.** An area efficient design for integer transform in H.264/AVC FRext / Y.K. Lin, Ying-Ze Liao and Tian-Sheuan Chang // The 17th VLSI Design/CAD Symposium, 2006.
 21. **Schuster, G.M.** Rate-Distortion Based Video Compression, Optimal Video Frame Compression, and Object Boundary Encoding / G.M. Schuster and A. Katsaggelos – Kluwer Academic Publisher, 1997.
 22. **Yeo, B.** Volume rendering of DCT-based compressed 3D scalar data / B. Yeo and B. Liu // IEEE Transactions on Visualization and Computer Graphics. – 1995. – P. 29-43.
 23. **Lee, M.** Quantization of 3D-DCT coefficients and scan order for video compression / M. Lee, R.K.W. Chan and D.A. Adjeroh // Journal of visual communication and image representation. – 1997. – Vol. 8(4). – P. 405-422.
 24. YUV Video Sequences, <http://trace.eas.asu.edu/yuv>.
 25. JM Video Codec, <http://iphome.hhi.de/suehring/tml>.
 26. MPEG-2 Video Codec, <http://mpeg.org/mpeg/mssg>.
 2. **Advanced video coding for generic audiovisual services.** ITU-T Recommendation H.264 and ISO/IEC 14496-10 (AVC), 2009.
 3. **ISO/IEC 13818 (MPEG-2):** Generic coding of moving pictures and associated audio information, 1994.
 4. **ITU-T and ISO/IEC JTC1.** «Digital Compression and coding of continuous-tone still images», ISO/IEC 10918-1, ITU-T Recommendation T.81 (JPEG), 1992.
 5. **Xiangyang Ji.** Concealment of Whole-Picture Loss in Hierarchical B-Picture Scalable Video Coding / Xiangyang Ji, Debin Zhao and Wen Gao // IEEE Transactions on Multimedia, -vol. 11, -P.11-22, 2008.
 6. **Z. Wang.** Image quality assessment: From error visibility to structural similarity / Z. Wang, Alan C. Bovik, Hamid R. Sheikh and Eero P. Simoncelli // IEEE Transactions on Image Processing, -vol. 13, -№. 4, -P. 600-612, 2004.
 7. **J. J. Koivusaari.** Simplified three-dimensional discrete cosine transform based video codec / J. J. Koivusaari and J. H. Takala // SPIE Proceedings in Multimedia on Mobile Devices, -vol. 5684, -P. 11–21, 2005.
 8. **Yui-Lam Chan.** Variable Temporal-Length 3-D Discrete Cosine Transform Coding / Yui-Lam Chan and Wan-Chi Siu // IEEE Transactions on Image Processing, Vol. 6, No. 5, 1997.
 9. **M. Servais.** Video Compression using the Three Dimensional Discrete Cosine Transform / M. Servais and Gerhard De Jager // Proceedings of the 1997 South African Symposium on Communications and Signal Processing, 1997.
 10. **R.K.W. Chan.** 3D-DCT Quantization as a Compression Technique for Video Sequences / R.K.W. Chan and M.C. Lee // Proceedings of the 1997 International Conference on Virtual Systems and MultiMedia, 1997.
 11. **B. Furht.** An Adaptive Three-Dimensional DCT Compression Based on Motion Analysis / B. Furht, Ken Gustafson, Hesong Huang and Oge Marques // Proceedings of the 2003 ACM symposium on Applied computing, 2003.
 12. **N. Krasilnikov.** Digital image processing – Moscow: «Vuzovskaya kniga» publisher, 2001. – 320 p. – (in Russian).
 13. **T. Mekky.** On the computation of the 3-D DCT / T. Mekky, S. Boussakta and M. Darnell // IEEE International Conference on Electronics, Circuits and Systems, vol. 3, pp. 1141 – 1143, 2003.
 14. **S.C. Chan.** Fast algorithms for computing the discrete cosine transform / S.C. Chan and K.L. Ho // IEEE Transactions on Circuits Systems II: Analog Digital Signal Process, vol. 39, pp.185-190, 1992.
 15. **M.A Chichyeva.** Image Block Coding Based on New Algorithms of Shortlength DCT with Minimal Multiplicative Complexity/ M.A Chichyeva, and V.M. Chernov // F.Solina, A.Leonardis (Eds).
 16. **T. Wiegand.** H.26L Test Model Long-Term Number 9 (TML-9) draft0 // ITU-T Q.6/SG16 Document TML9.doc (VCEG-N83d1), 2001.
 17. **A. Hallapuro.** Low complexity transform and quantization – Part 1: Basic Implementation / A. Hallapuro and M. Karczewicz // JVT document JVT-B038, 2002.
 18. **H.S. Malvar.** Low-complexity transform and quantization in H.264/AVC / H.S. Malvar, A. Hallapuro, M. Karczewicz and L. Kerofsky // IEEE Transactions on Circuits and Systems for Video Technology, Vol. 13(7), pp. 598-603, 2003.
 19. **S. Gordon.** Simplified Use of 8x8 Transform // Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, doc. JVT-I022, 2003.
 20. **Y.K. Lin.** An area efficient design for integer transform in H.264/AVC FRext / Y.K. Lin, Ying-Ze Liao and Tian-Sheuan Chang // The 17th VLSI Design/CAD Symposium, 2006.

References

1. **Zhihai, He.** Power-rate-distortion analysis for wireless video communication under energy constraints / Zhihai He, Y. Liang, L. Chen, I. Ahmad and Dapeng Wu // IEEE Transactions on Circuits and Systems for Video Technology, -vol.15, -P.645-658, 2005.

21. **G.M. Schuster**, Rate-Distortion Based Video Compression, Optimal Video Frame Compression, and Object Boundary Encoding / G.M. Schuster and A. Katsaggelos, Kluwer Academic Publisher, 1997.
22. **B. Yeo**, Volume rendering of DCT-based compressed 3D scalar data / B. Yeo and B. Liu // IEEE Transactions on Visualization and Computer Graphics, pp.29-43, 1995.
23. **M. Lee**, Quantization of 3D-DCT coefficients and scan order for video compression / M. Lee, R.K.W. Chan and D.A. Adjeroh, Journal of visual communication and image representation, Vol.8 (4), pp.405-422, 1997.
24. **YUV Video Sequences**, <http://trace.eas.asu.edu/yuv>.
25. **JM Video Codec**, <http://iphome.hhi.de/suehring/tml>.
26. **MPEG-2 Video Codec**, <http://mpeg.org/mpeg/mssg>.

VIDEO COMPRESSION BASED ON TREE-DIMENSIONAL PSEUDO COSINE TRANSFORM FOR VIDEO SURVEILLANCE SYSTEMS

E. A. Belyaev¹, T. M. Sukhov², N. N. Shostackiy³

¹ *Saint-Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences,*

² *Baltic State Technical University,*

³ *JSC «TIRS»*

Abstract

This paper discusses video compression techniques based on three-dimensional discrete cosine transformation for video surveillance systems. Multiplication free three-dimensional discrete pseudo cosine transformation, motion analyzer and low-complexity temporal filtration method are proposed. Comparison results which show the practical efficiency of the proposed algorithm in comparison with H.264/AVC и MPEG-2 standards are presented.

Key words: video compression, three-dimensional discrete cosine transform, temporal filtration.

Сведения об авторах



Беляев Евгений Александрович, 1981 года рождения. В 2005 году с отличием окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения (СПбГУАП) по специальности «Автоматизированные системы обработки информации и управления». Кандидат технических наук (2009 год), работает научным сотрудником в лаборатории информационных технологий в системном анализе и моделировании Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИ РАН) и ассистентом кафедры информационно-сетевых технологий СПбГУАП. В списке научных работ Е.А. Беляева 28 статей и 2 патента. Страница в интернете: <http://www.eugeny-belyaev.narod.ru>. E-mail: e_belyaev@mail.ru. Область научных интересов: сжатие и передача видеoinформации.

Eugeny Alexandrovich Belyaev (b. 1981) graduated with honours (2005) from the St.-Petersburg State University of Aerospace Instrumentation (SUAI), majoring in Automation and Control. He received his Candidate (2009) degree in system analysis, control and information processing from SUAI. He is the Research Scientist in the Laboratory of Information technologies in systems analysis and modeling of St.-Petersburg Institute for Informatics and Automation (SPIIRAS) and Assistant Professor in St.-Petersburg State University of Aerospace Instrumentation. He is co-author of 28 scientific papers and 2 patents. Homepage: http://www.eugeny-belyaev.narod.ru/ebelyaev_eng.htm. E-mail: e_belyaev@mail.ru. His current research interests include real-time video compression and transmission, video source rate-control, scalable video coding, motion estimation and arithmetic encoding.



Сухов Тимофей Михайлович, 1985 года рождения, в 2007 году окончил Балтийский государственный технический университет (БГТУ) «ВоенМех» им. Д.Ф. Устинова по специальности «Информационные системы и технологии в телекоммуникациях», аспирант кафедры радиоэлектронных систем управления БГТУ «ВоенМех». E-mail: tsuhov@mail.ru. Область научных интересов: сжатие изображений и видео, обработка изображений, цветовые преобразования.

Timofey Mihajlovich Sukhov (b. 1985) graduated from the Baltic State Technical University (BSTU) "VoenMeh" of D.F.Ustinova (BSTU), majoring in Information systems and technologies in telecommunications. He is the post-graduate student of radio-electronic control systems department of BSTU. E-mail: tsuhov@mail.ru. His current research interests include compression of images and video, processing of images and color transformations.



Шостацкий Николай Николаевич, 1938 года рождения. В 1967 году с отличием окончил Ленинградский горный институт (ныне Санкт - Петербургский государственный горный институт) по специальности «Радиотехника». Кандидат технических наук (1975 год), старший научный сотрудник. Работает научным консультантом в научно - производственной фирме «ТИРС» (г. Санкт – Петербург). В списке научных трудов Н.Н. Шостацкого 82 статьи, одна монография и 24 авторских свидетельства. E-mail: *shostacki@yandex.ru*.

Область научных интересов: цифровое кодирование и передача телевизионных изображений по каналам связи, цифровая обработка изображений.

Nikolay Nikolaevich Shostackiy (b. 1938) graduated with honours (1967) from the Leningrad Mining Institute (presently, Saint-Petersburg State Mining Institute), majoring in Radio Engineering. He received his Candidate of Science in engineering (1975). He is Senior Research scientist and Scientific Adviser in JSC «TIRS» (Saint-Petersburg). He is co-author of 82 scientific papers, 1 monograph, and 24 inventions. E-mail: *shostacki@yandex.ru*. His current research interests include digital coding and transmission of television images and digital signal processing.

Поступила в редакцию 16 марта 2010 г.