

И. С. РУБИНА, А. Ю. ТРОПЧЕНКО

ИССЛЕДОВАНИЕ АЛГОРИТМОВ КОДИРОВАНИЯ ПРЕОБРАЗОВАНИЕМ В ЗАДАЧАХ СЖАТИЯ КАДРОВ ВИДЕОПОСЛЕДОВАТЕЛЬНОСТИ

Предложен быстрый алгоритм кодирования видеоданных преобразованием при сжатии опорных и остаточных кадров, основанный на трехмерном преобразовании Хартли с фиксированным и переменным размером матрицы преобразования.

Ключевые слова: кодирование преобразованием, преобразование Хартли, трехмерный алгоритм, матрица переменного размера, пространственная модель.

Введение. Алгоритмы кодирования преобразованием широко используются для сжатия изображений, а также видеопоследовательностей. Этап кодирования преобразованием является частью пространственной модели видеокомпрессора [1]. Исходными данными для этого этапа являются отсчеты кадров видеопоследовательности. Обрабатываемые кадры подразделяются на:

1) опорные, или intra-кадры [1], не обрабатываемые блоком компенсации движения. Чем больше таких кадров выделяется в процессе сжатия видеопоследовательности, тем выше качество воспроизводимого видеоряда и меньше степень сжатия;

2) остаточные, представляющие собой разность исходного кадра и кадра-прогноза, полученного в процессе компенсации движения. Такие кадры определяют погрешность временной модели. В результате выполнения кодирования формируются коэффициенты разложения сигнала по базису, состоящему из функций, локализованных по частоте, а также в пространстве. Полученные коэффициенты преобразования поступают на вход квантователя и переупорядочиваются. Процесс сжатия завершается поэлементной обработкой потока энтропийным кодером.

Алгоритмы кодирования преобразованием подразделяются на блочные и глобальные, когда преобразованию подвергается все изображение [2]. В настоящей работе анализируется блочный тип преобразования.

Блочные преобразования выполняются над квадратными блоками опорного или остаточного кадра и после ряда операций формируют блоки коэффициентов таких же размеров. Элементами таких блоков служат сэмплы. Любой блок кадра можно восстановить с помощью линейной комбинации $N \times N$ базисных шаблонов [1], умножаемых на соответствующие весовые множители (коэффициенты преобразования). Достоинством блочных преобразований являются низкие требования к объему памяти. Среди недостатков можно выделить возможность возникновения артефактов на стыках блоков.

При кодировании на основе блоков используются ортогональные преобразования в различных функциональных базисах, например, дискретное косинусное преобразование (ДКП) и преобразование Хартли. Достоинством преобразования Хартли является то, что в отличие от преобразования Фурье оно вещественное. Преобразование Хартли определяется как

$$H(s) = \int_{-\infty}^{\infty} f(x) \text{cas} 2\pi s x dx, \quad (1)$$

где $\text{cas} 2\pi s x = \sin 2\pi s x + \cos 2\pi s x$ [3].

Схемы алгоритмов. В ходе исследования был проведен анализ следующих алгоритмов преобразования.

1. Двумерное дискретное преобразование Хартли (2D-ДПХ), определяемое формулой:

$$H(u, v) = \frac{1}{NM} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} f(i, j) \operatorname{cas} \left(\frac{2\pi ui}{N} + \frac{2\pi vj}{M} \right), \quad (2)$$

где $N \times M$ — размеры матрицы преобразования, а $f(i, j)$ — значение яркости кадра в точке с координатами (i, j) .

2. Двумерное дискретное косинусное преобразование (2D-ДКП), определяемое формулой:

$$Y(u, v) = C_u C_v \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} f(i, j) \cos \frac{xu(2i+1)}{2N} \cos \frac{y\pi(2j+1)}{2M}, \quad (3)$$

где $C_u = \sqrt{\frac{1}{N}}$ ($u=0$), $C_u = \sqrt{\frac{2}{N}}$ ($u>0$), а $Y(i, j)$ — коэффициенты разложения.

3. Трехмерное дискретное преобразование Хартли (3D-ДПХ), определяемое в соответствии с выражением:

$$X(u, v, t) = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \sum_{k=0}^{P-1} \operatorname{cas} \left(\frac{2\pi}{N} iu + \frac{2\pi}{M} jv + \frac{2\pi}{P} kt \right) f(i, j, k), \quad (4)$$

где P — число кадров, индекс t определяет номер двумерной страницы в трехмерном блоке разложения, а $f(i, j, k)$ — значение яркости отсчета в точке с координатами (i, j) k -го кадра.

Рассмотренные двумерные алгоритмы преобразования реализуются строчно-столбцовым методом. В матричном виде это может быть записано следующим образом:

$$F_N = [E_N X_N] E_N, \quad (5)$$

где E_N — матрица ядра дискретного ортогонального преобразования.

Дискретное многомерное преобразование Хартли реализовывалось в соответствии с алгоритмом, описанным в работе [4]. Такой подход позволяет существенно снизить вычислительные затраты.

4. Быстрое трехмерное дискретное преобразование Хартли с фиксированным размером матрицы преобразования (3D-БПХФ). Разработанный быстрый алгоритм основан на вычислении коэффициентов меньших 3D-блоков через коэффициенты более крупных.

Данный алгоритм быстрого преобразования получен в результате декомпозиции исходной суммы основного соотношения на восемь частичных сумм, для каждой из которых i, j, k либо четные, либо нечетные, и применения тригонометрических преобразований функций:

$$\left. \begin{aligned} x_0 &= X'_{000}(u, v, t) \\ x_1 &= X'_{001}(u, v, t) \operatorname{cas} \frac{2\pi t}{P}, \\ x_2 &= X'_{010}(u, v, t) \operatorname{cas} \frac{2\pi v}{M}, \\ x_3 &= X'_{011}(u, v, t) \operatorname{cas} 2\pi \left(\frac{v}{M} + \frac{t}{P} \right), \\ x_4 &= X'_{100}(u, v, t) \operatorname{cas} \frac{2\pi u}{N}, \\ x_5 &= X'_{101}(u, v, t) \operatorname{cas} 2\pi \left(\frac{u}{N} + \frac{t}{P} \right), \\ x_6 &= X'_{110}(u, v, t) \operatorname{cas} 2\pi \left(\frac{u}{N} + \frac{v}{M} \right), \\ x_7 &= X'_{111}(u, v, t) \operatorname{cas} 2\pi \left(\frac{u}{N} + \frac{v}{M} + \frac{t}{P} \right), \end{aligned} \right\} \quad (6)$$

где

$$X'_{abc}(u, v, t) = \sum_{i=0}^{N/2-1} \sum_{j=0}^{M/2-1} \sum_{k=0}^{P/2-1} X(2i+a, 2j+b, 2k+c) \operatorname{cas} \left(2\pi \left(\frac{2iu}{N} + \frac{2jv}{M} + \frac{2kt}{P} \right) \right),$$

abc — трехбитный двоичный код, определяющий номер формируемой суммы.

Блок промежуточной размерности формируется на основе вычисляемых частичных сумм по формуле:

$$X(u', v', t') = \sum_{r=0}^7 (\pm x_r), \quad (7)$$

где u', v', t' — размерность вычисляемого промежуточного блока, а r — номер частичной суммы. Промежуточный блок может иметь размерность, уменьшаемую вдвое хотя бы по одной из координат. Если размер блока уменьшается по одной или по всем координатам, то знак суммы считается отрицательным. При уменьшении размеров по двум координатам знак не изменяется.

5. Быстрое трехмерное дискретное преобразование Хартли с переменным размером матрицы преобразования (3D-БПХП). Выбор размера матрицы преобразования определяется на этапе компенсации движения по следующему алгоритму [5]:

- все изображение делится на блоки равного размера;
- для каждого блока выполняется проверка условия: если величина ошибки (соответствие) обрабатываемого блока с наиболее схожим блоком ссылочного кадра выше некоторой пороговой, то блок делится на четыре промежуточных фрагмента меньшего размера;
- если достигнуто заданное число промежуточных блоков или получена требуемая точность совпадения, процесс для исходного блока останавливается.

Экспериментальные результаты. Для анализа описанных ранее алгоритмов кодирования преобразованием были выбраны стандартные тестовые последовательности группы MPEG: „Теннис“, „Бригадир“ и „Береговая охрана“ [6].

Для количественного сравнения эффективности рассматриваемых преобразований были использованы следующие характеристики:

1) пиковое соотношение сигнал/шум (PSNR, peak signal to noise ratio) (рис. 1), вычисляемое в соответствии с формулой:

$$\text{PSNR} = 10 \log_{10} \frac{(2^q - 1)^2}{\text{MSE}}, \quad (8)$$

где $q=8$ — разрядность цветовой схемы, а MSE — среднеквадратичное отклонение яркости исходного изображения от восстановленного, определяемое формулой:

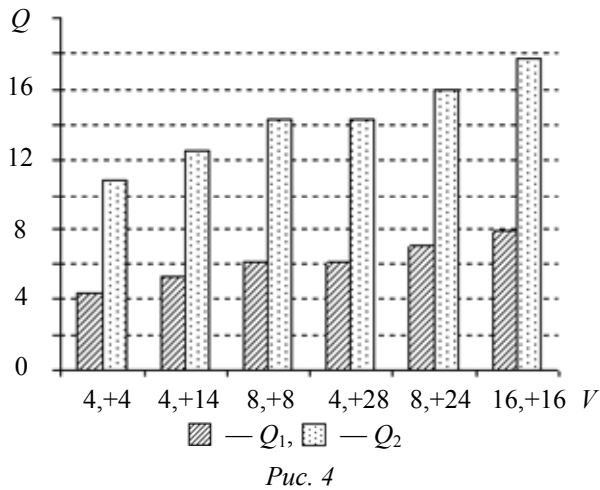
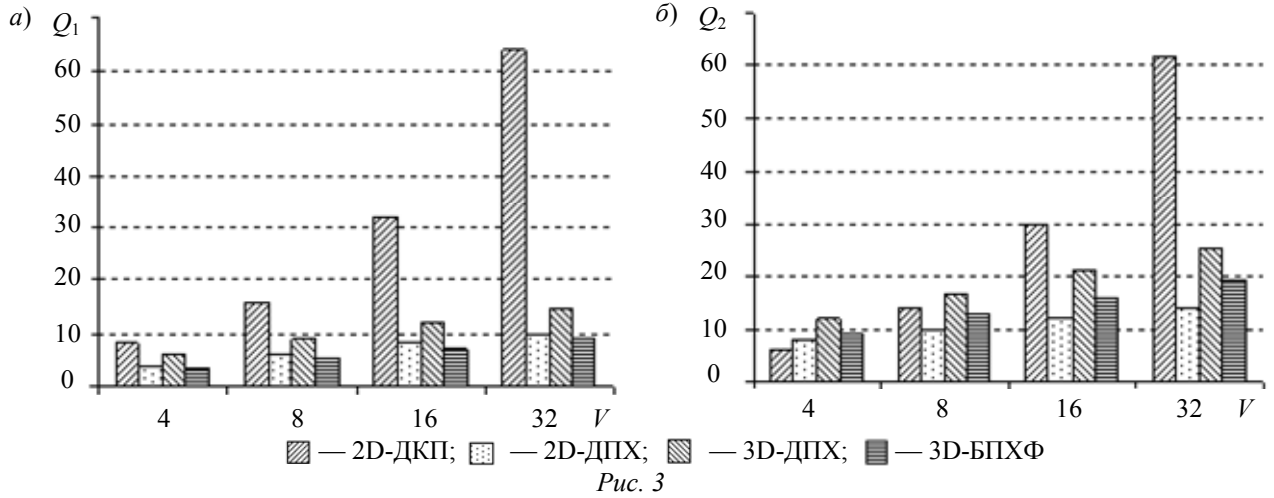
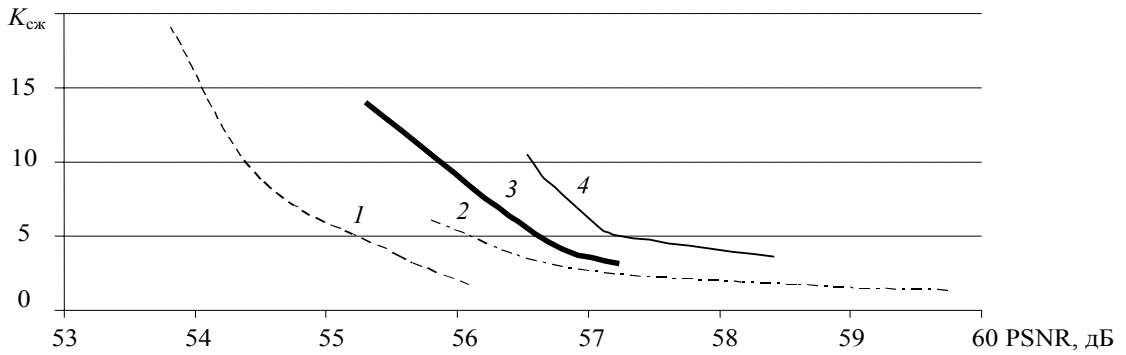
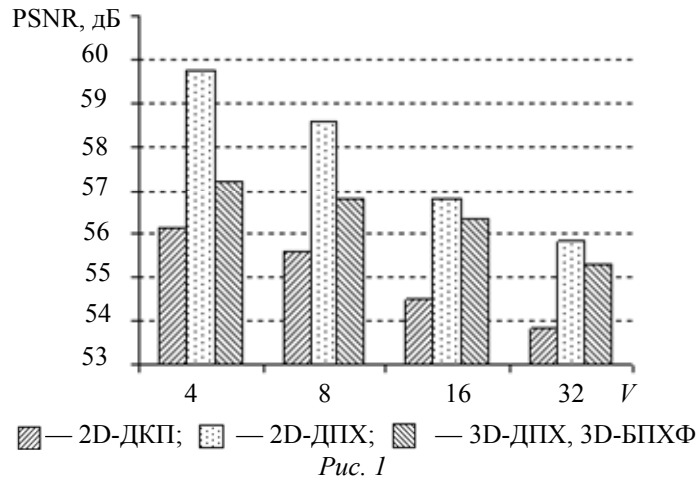
$$\text{MSE} = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \|f(i, j) - \hat{f}(i, j)\|^2, \quad (9)$$

где $f(i, j)$ и $\hat{f}(i, j)$ — яркость соответствующих пикселей исходного и восстановленного кадра;

2) характеристика, выражающая зависимость искажения сигнала (PSNR) от степени его сжатия $K_{\text{сж}}$ (рис. 2, 1 — 2D-ДКП; 2 — 2D-ДПХ; 3 — 3D-ДПХ, 3D-БПХФ; 4 — 3D-БПХП);

3) вычислительная сложность алгоритма, измеряемая числом операций умножения Q_1 и сложения Q_2 на пиксел (рис. 3, a и b соответственно). На рис. 4 представлены зависимости числа операций Q от размера блока V с учетом возможного приращения его размерности.

Все эксперименты проводились для 256 уровней квантования.



Заключение. В ходе анализа было установлено следующее:

1) алгоритм 2D-ДПХ превзошел 2D-ДКП по качеству восстановленной видеопоследовательности на 5 %, но уступил ему в степени сжатия;

2) алгоритм 3D-ДПХ позволил вдвое увеличить степень сжатия видеопоследовательности при незначительном ухудшении ее качества при восстановлении. Это объясняется тем, что алгоритм выполняет преобразование не только в пространстве, но и во времени, устраняя соответствующую избыточность;

3) алгоритм 3D-БПХФ позволил на 30 % сократить число операций сложения/умножения на пиксел кадра видеопоследовательности за счет иерархического расчета коэффициентов преобразования;

4) алгоритм 3D-БПХП обеспечил повышение качества восстановленной видеопоследовательности на 2 %, а также степень ее сжатия на 1,5 %, что объясняется использованием адаптивно выбираемого размера матрицы преобразования для областей с мелкими деталями и для областей фона соответственно.

Таким образом, при использовании предложенных алгоритмов получено более высокое качество восстановленной видеопоследовательности и большее значение коэффициента сжатия при меньших вычислительных затратах за счет применения алгоритмов быстрого преобразования и переменного размера матрицы преобразования.

СПИСОК ЛИТЕРАТУРЫ

1. Ричардсон Я. Видеокодирование. H.264 и MPEG-4 — стандарты нового поколения. М.: Техносфера, 2005.
2. Ватолин Д., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. М.: ДИАЛОГ-МИФИ, 2003.
3. Брейсуэлл Р. Преобразование Хартли. М.: Мир, 1990.
4. Yonghong Z., Guoan B., Abdul R. L. New algorithms for multidimensional discrete Hartley transform // Signal Processing. 2002. Vol. 82. P. 1086—1095.
5. Ribas-Corbera J., Neuhoff D. L. On the optimal block size for block-based, motion compensated video coders // SPIE Proc. of Visual Communications and Image Processing. 1997. Vol. 3024. P. 1132—1143.
6. Рубина И. С., Тропченко А. Ю. Исследование алгоритмов выбора опорных пикселей в задачах выделения сегментов кадра видеопоследовательности // Изв. вузов. Приборостроение. 2012. Т. 55, № 1. С. 9—14.

Сведения об авторах

Ирина Семеновна Рубина

— аспирант; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра вычислительной техники; E-mail: rubren@mail.ru

Александр Ювенальевич Тропченко

— д-р техн. наук, профессор; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра вычислительной техники; E-mail: tau@d1.ifmo.ru

Рекомендована кафедрой
вычислительной техники

Поступила в редакцию
08.02.12 г.