

**Компенсация движения для видеокодирования
на основе гексагональных блоков**

С.В. Умняшкин, И.О. Шаронов

Национальный исследовательский университет «МИЭТ»

**Hexagonal Block Motion Compensation
for Video Compression**

S.V. Umnyashkin, I.O. Sharonov

National Research University of Electronic Technology, Moscow

Исследован алгоритм компенсации движения на основе блоков шестиугольной формы для видеокомпрессии. Полученное после компенсации движения разностное изображение сжималось с помощью алгоритма SPIHT, основанного на дискретном вейвлет-преобразовании. Экспериментально подтверждено, что использование шестиугольной формы блоков повышает визуальное качество обработанной видеопоследовательности, а применение многомодельной статистики в послойной схеме кодирования позволяет уменьшить сложность разработанного алгоритма.

Ключевые слова: шестиугольные блоки; дискретное вейвлет-преобразование; компенсация движения; видеокомпрессия.

The algorithm of the layered motion compensation based on the hexagonal shape of the blocks has been improved. The difference image, produced after the motion compensation has been compressed by the SPIHT algorithm based on a discrete wavelet transform. It has been experimentally proved that using hexagonal blocks improves the visual quality of processed video sequence and statistical multi-modeling in layered encoding has enabled to decrease the computational complexity of the video compression scheme proposed.

Keywords: hexagonal blocks; discrete wavelet transform; motion compensation; video compression.

Введение. Модуль блочной компенсации движения, как правило, является одним из основных в разных методах видеокомпрессии, так как именно он обеспечивает устранение временной избыточности между кадрами [1]. Первый кадр видеопоследовательности сжимается как статическое изображение, компенсация движения к нему не

применяется. Такой кадр называется ключевым, или *I*-кадром. Каждый следующий кадр сравнивается с предыдущим закодированным кадром поблочно для нахождения похожих областей на изображениях. С помощью векторов, которые характеризуют перемещения блоков, для текущего кадра строится изображение из блоков предыдущего кадра. Полученное прогнозируемое изображение затем вычитается из оригинального кадра, и в результате формируется разностное изображение, кодируемое далее кодером статических изображений. Векторы перемещения кодируются отдельно энтропийным кодером [2]. Кадры, полученные с помощью компенсации движения по одному предыдущему кадру, называются *P*-кадрами. Данный процесс повторяется многократно, пока все кадры последовательности не будут закодированы.

Традиционно для блочной компенсации движения применяются блоки прямоугольной (обычно квадратной) формы. В настоящей работе рассматривается реализация, использующая блоки шестиугольной формы. Оригинальная идея предложена в работах [3, 4]. Форма блоков и их размещение на кадре представлены на рис.1. Как видно, блок вкладывается в прямоугольную область размером 8×10 пикселей и покрывает собой 64 пикселя. Левые и правые граничные блоки имеют иную структуру.

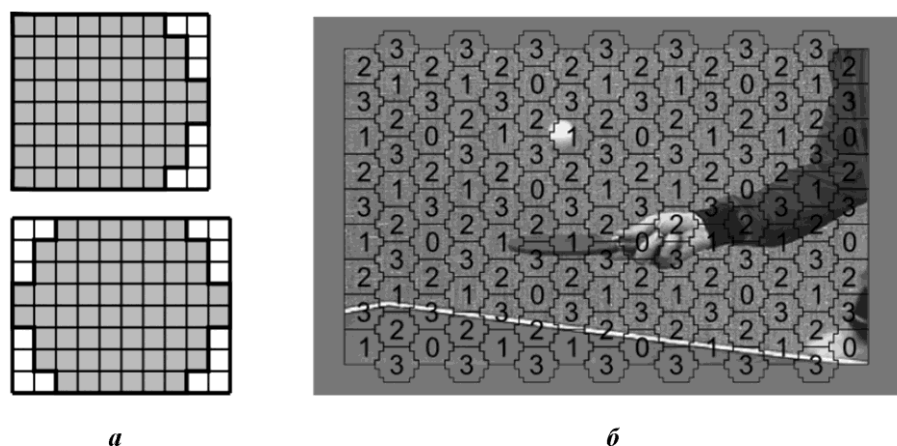


Рис.1. Шаблон блока шестиугольной формы (а) и схематическое размещение блоков на кадре по слоям (цифры соответствуют номерам слоя) (б)

Покрыть всю область прямоугольного кадра шестиугольными блоками не представляется возможным. Поэтому изображение сначала расширяется (экстраполируется), а затем блоки размещаются таким образом, чтобы покрыть всю активную часть кадра. Эксперименты показали, что экстраполяция кадра постоянным значением 127 оказывает лучшее влияние на качество восстановленного видеоряда, чем стандартное расширение значением последнего граничного пикселя [1]. При использовании экстраполяции кадра производительность кодека несколько снижается вследствие увеличения области поиска векторов перемещений граничных блоков в расширенной зоне. Для видеопоследовательностей разрешением 352×288 пикселей производительность снижается примерно на 5%.

Сегментация кадра на слои. Разобьем исходный кадр на слои – группы блоков, объединенных специальным образом (на рис.1 представлены группы с цифрами 0, 1, 2 и 3). Такое разбиение на четыре слоя представляет собой определенную модификацию трехслойной схемы, рассмотренной в [3].

Нулевой слой предназначен для оценивания глобального перемещения в кадре. Так как результирующие векторы сразу кодируются, дополнительные затраты на пересчет

векторов не требуются. Блоки для нулевого слоя фактически представляют собой треть блоков первого слоя и формируют далее прогноз векторов первого слоя по векторам блоков нулевого слоя (см. рис.1). Для поиска векторов используется шаблон №1 [3, 5] с областью поиска наибольшего размера (рис.2). В зонах 1 и 2 поиск осуществляется в два этапа: сначала ищется наилучший вектор на отмеченных цветом позициях, затем для него проверяются ближайшие соседние позиции.

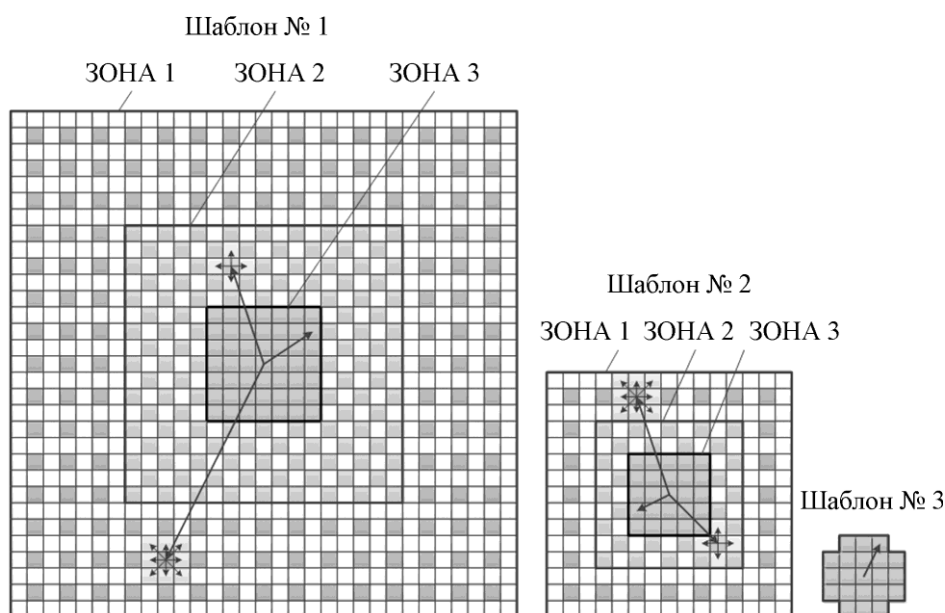


Рис.2. Шаблоны поиска векторов перемещений

Векторы перемещений блоков первого слоя ищутся на основе найденных векторов нулевого слоя: область поиска смещается на величину прогнозного вектора, затем производится поиск по шаблону либо №1, либо №2 (см. рис.2) [3, 5]. Прогнозный вектор находится путем медианной фильтрации векторов окружающих блоков нулевого слоя.

Область поиска для каждого блока смещается на вектор, который получается путем медианной фильтрации уже найденных векторов перемещений соседних блоков первого слоя. Используются шаблоны поиска №2 или №3 (см. рис.2).

На третьем слое используется шаблон №3 с областью поиска наименьшего размера, а прогнозный вектор находится в результате медианной фильтрации векторов перемещений уже всех соседних блоков.

Многомодельность. Чем больше векторов участвуют в прогнозе, тем точнее можно предсказать вектор перемещения. Поскольку с увеличением номера слоя увеличивается и количество окружающих векторов для прогноза, то размер шаблонов, определяющих площадь области поиска в пикселях, уменьшается для сокращения объема вычислений. Следует ожидать, что распределение найденных векторов будет отличаться на разных слоях. Более того, среднеквадратическое отклонение компонент векторов, которые участвуют в прогнозе, характеризует и разброс погрешности этого прогноза.

Данные предпосылки определяют необходимость использования нескольких статистических моделей для каждого слоя. На рис.3 показано, как соотносятся модели, используемые для арифметического кодирования векторов перемещений, и шаблоны поиска. Модели выбираются исходя из среднеквадратического отклонения компонент

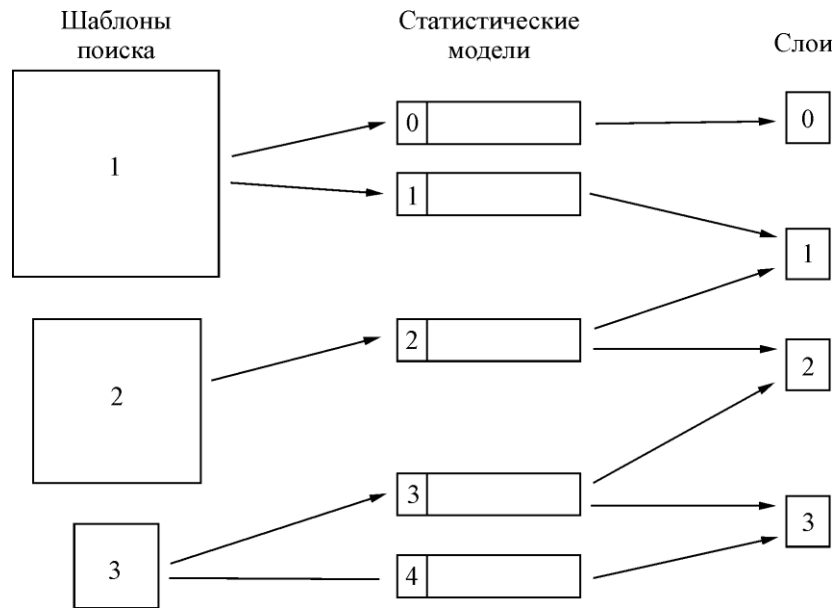


Рис.3. Статистические модели и шаблоны поиска для разных слоев

векторов, используемых при прогнозе. Каждая модель имеет свой шаблон поиска. Таким образом, 1-й и 2-й слои в некоторых случаях могут производить поиск по шаблонам меньшего размера. Каждый слой, кроме нулевого, имеет по две модели. Все модели являются динамически пополняемыми, т.е. в них представлены не все возможные векторы перемещений (которые рассматриваются как некоторые символы алфавита модели), а только те, которые уже встречались. Если в модели еще нет требуемого символа, то при первом его появлении сначала кодируется маркер нового символа [6].

Поиск векторов перемещений. Для каждого i -го блока поиск векторов перемещений осуществляется как поиск минимума функции

$$J_i(\mathbf{r}) = D_i(\mathbf{r}) + \lambda_{mv} R_i(\mathbf{r}), \quad (1)$$

где \mathbf{r} – вектор перемещения с координатами r_x и r_y ; D_i – искажение, вносимое от замены блока текущего кадра, смещенным на вектор \mathbf{r} , блоком из предыдущего декодированного кадра; $R_i(\mathbf{r}) = -\log_2 v(\mathbf{r})$ – оценка битовых затрат на кодирование вектора перемещения \mathbf{r} , которая вычисляется по частоте его появления $v(\mathbf{r})$ в выходном потоке среди других векторов; λ_{mv} – весовой коэффициент Лагранжа, отвечающий за баланс между качеством восстановленного видео и степенью сжатия векторов перемещений.

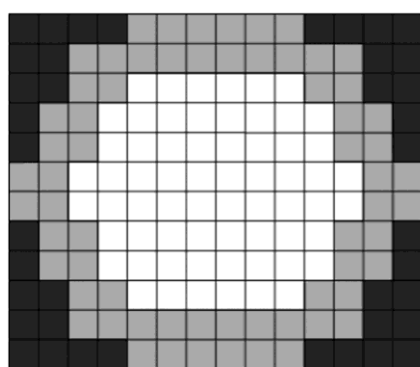
Все векторы перемещения сжимаются алгоритмом арифметического кодирования [2] с накапливающейся статистикой, на основе которой вычисляются битовые затраты. Мера искажения D_i рассчитывается по формуле среднеквадратического отклонения.

Использование масок. Для того чтобы избежать появления на разностном изображении резких перепадов яркости на границах блоков, приводящих к снижению эффективности сжатия, в видеокодеке используется блочная компенсация с перекрытием [1].

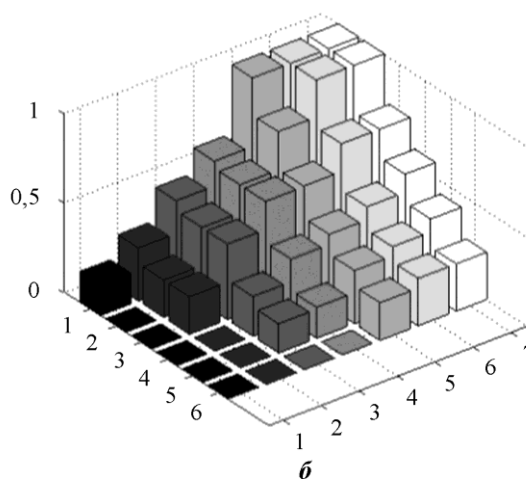
В настоящей работе в качестве оконной функции (или просто «маски») использовалась специальная матрица, построенная на основе двух коэффициентов a и b . Матрица имеет следующий вид [3, 4]:

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & \frac{a}{1+2a} & \frac{a}{1+a} & \frac{a}{1+a} \\ 0 & 0 & \frac{ab}{1+ab+a} & \frac{a}{1+a+b} & \frac{a}{1+2b} & \frac{a}{1+a+b} & \frac{a}{1+b} \\ 0 & 0 & \frac{a}{1+ab+b} & \frac{b}{1+a+b} & \frac{1}{1+2b} & \frac{1}{1+a+b} & \frac{1}{1+b} \\ 0 & \frac{ab}{1+ab} & \frac{b}{1+b} & \frac{1}{1+b} & \frac{1}{1+2a} & \frac{1}{1+a} & \frac{1}{1+a} \\ 0 & \frac{1}{1+2a} & \frac{1}{1+b} & \frac{1}{1+b} & \frac{1}{1+ab} & 1 & 1 \\ a & \frac{b}{1+2a} & \frac{1}{1+b} & \frac{1}{1+b} & 1 & 1 & 1 \\ 1+a+b & 1+2b & 1+a+b & 1+ab+a & 1 & 1 & 1 \end{pmatrix}.$$

Маска, которая обеспечивает минимальную среднеквадратическую ошибку после сжатия, считается оптимальной и имеет эмпирически найденные коэффициенты: $a = 0,36$ и $b = 0,73$ (рис.4).



а



б

Рис.4. Границы блочной маски (а) и левая нижняя четверть оптимальной маски (б)

Послойное кодирование. После обработки p -го слоя кодирование векторов перемещений блоков может быть остановлено, если достигается минимум функции:

$$J_l(p) = D_p + \lambda_l \sum_{k \leq p} R_k, \quad (2)$$

где D_p – среднеквадратическое отклонение между оригинальным и полученным прогнозным кадром; λ_l – параметр для управления качеством/степенью сжатия; R_k – битовые затраты на кодирование k -го слоя.

Недостающие векторы блоков с отброшенных слоев заменяются векторами прогнозов, что позволяет экономить ресурсы на поиск и снижать битовые затраты.

Кодирование разностного изображения. Разница между оригинальным и прогнозным кадрами (разностное изображение) кодируется с помощью алгоритма SPIHT [7], основанного на дискретном вейвлет-преобразовании. Кодирование останавливается, если для i -й битовой плоскости вейвлет-коэффициентов закодированного разностного кадра достигается минимум RD -функции:

$$J_s(i) = D_i + \lambda_s R_i, \quad (3)$$

в котором отклонение D_i определяется как среднеквадратическая ошибка от отбрасывания битовых плоскостей, а R_i – количество битов, затраченных на кодирование i -й битовой плоскости и всех предшествующих. Как и для кодирования векторов перемещений, параметр λ_s задается извне алгоритма кодирования и отвечает за баланс между степенью сжатия и качеством.

Управление RD-зависимостью. Как показывают уравнения (1)–(3), регулировать зависимость между качеством восстановленного видеоряда и степенью его сжатия можно с помощью свободных коэффициентов λ_{mv} , λ_l и λ_s . Чем больше соответствующие коэффициенты, тем степень сжатия будет больше, а качество – меньше.

Поиск векторов перемещений основывается на сравнении опорного (предыдущего декодированного) кадра с оригинальным кодируемым. Точность нахождения вектора зависит от коэффициента λ_{mv} и качества восстановления опорного кадра. Последнее, в свою очередь, зависит от наличия отброшенных слоев, т.е. от коэффициента λ_l , и от степени сжатия опорного кадра, которое управляется коэффициентом λ_s .

Таким образом, для управления зависимостью между битрейтом видео и его качеством достаточно оперировать описанными коэффициентами. Логично ввести один коэффициент λ , от которого зависят все свободные коэффициенты в (1)–(3). Простейшая модель такой зависимости линейная. Экспериментально получено, что баланс между производительностью, степенью сжатия и качеством восстановленной видеопоследовательности достигается в этом случае при $\lambda_s = \lambda$, $\lambda_{mv} = \lambda$, $\lambda_l = 20\lambda$. Для построения RD-кривых (см. рис.4 и 5) изменялся только коэффициент λ .

Производительность. Для уменьшения времени поиска векторов в предлагаемой реализации используется большее (по сравнению с рассмотренными в [3]) количество моделей арифметического кодера: для каждого слоя определены по две «смежные» модели. Только у нулевого слоя одна модель. Одна из двух моделей слоя выбирается в зависимости от величины дисперсии векторов блоков, по которым строился прогноз, в результате сравнения с заданным для каждого слоя пороговым значением. Всего имеем пять моделей. При переходе от одной модели к другой шаблоны поиска также переключаются (см. рис.3). Это позволяет уменьшить область поиска при малой дисперсии прогнозных векторов и этим увеличить производительность в итоге примерно в 1,8 раза.

Механизм послойного кодирования с отбрасыванием слоев позволяет на низких битрейтах дополнительно увеличить производительность еще на 5–17%, при этом битовые затраты на кодирование отброшенных слоев отсутствуют.

Практические результаты. Проведем сравнение видеокodeка с межкадровой блочной компенсацией на основе шестиугольных блоков и аналогичного codeка с квадратными блоками [8]. Для сравнения видеокodeков рассматривалась объективная метрика визуального качества [1] – величина PSNR – при одинаковых битовых затратах на кодирование видеопоследовательностей, выраженных в количестве битов на пиксель (bits per pixel, bpp):

$$R = \frac{B}{L \cdot M \cdot N} [\text{bpp}],$$

где B – количество битов выходного потока; L – общее количество кадров; M и N – пиксельная размерность кадра видеопоследовательности.

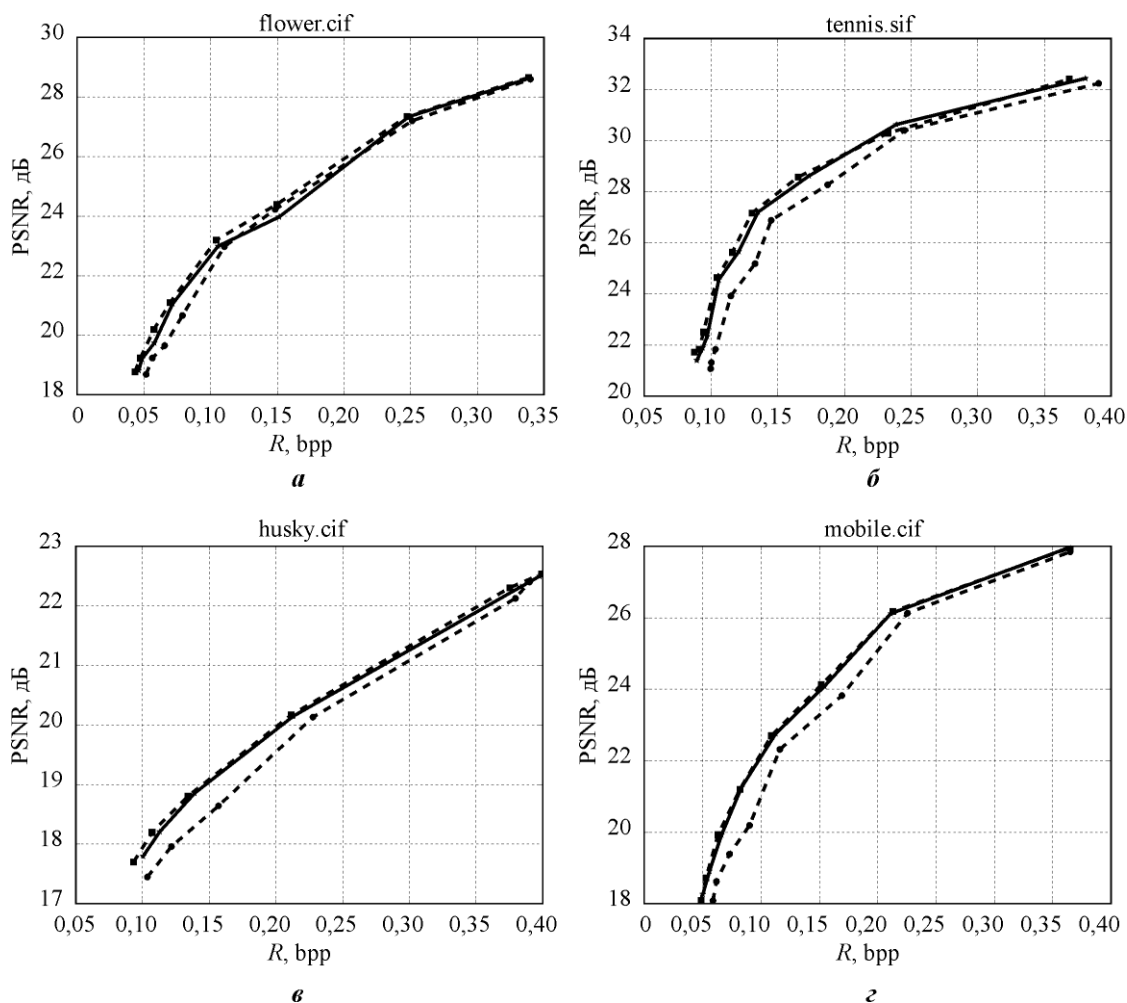


Рис.5. Сравнение шестиугольной и квадратной форм блоков на примере различных видеопоследовательностей: Flower Garden (а), Tennis (б), Husky (в), Mobile (г): ●—● квадратные блоки; ■—■ шестиугольные блоки (одна модель на слой); — шестиугольные блоки М (М – применение многомодельной схемы с послойным кодированием)

Величина PSNR определялась как

$$\text{PSNR} = \frac{1}{L} \sum_{k=1}^L 10 \lg \frac{255^2}{\text{MSE}(I_k, \tilde{I}_k)} \quad [\text{дБ}]$$

где L – общее количество кадров; k – порядковый номер кадра; MSE – среднеквадратическое отклонение; I_k – оригинальный кадр; \tilde{I}_k – восстановленный кадр.

Применение многомодельной схемы (см. рис.3) значительно повышает производительность кодека, но несколько понижает качество восстановленного видеоряда. Как показывают эксперименты (см. рис.5), снижение PSNR в большинстве случаев оказывается несущественным. Дополнительное применение послойного кодирования с отбрасыванием слоев слабо влияет на визуальное качество видео. Отбрасывание слоев начинает работать в среднем при $R \approx 0,1-0,15$ bpp и меньше.

Сравнения проводились с кодеком [5], основанным на квадратных блоках [8], с модифицированной многомодельностью, соответствующей рис.3. Результаты сравнения представлены также на рис.5. В обоих кодеках для обработки разнотных изображений применялось двумерное дискретное вейвлет-преобразование (коэффициенты биортого-

нальных фильтров взяты из работы [9]): для кадра размером cif (352×288 пикселей) выполнялось пять уровней декомпозиции, а для кадра sif (352×240) – четыре уровня.

Следует отметить, что визуальное восприятие глазом качества восстановленного видео для шестиугольных блоков оказывается более высоким, чем для квадратных блоков (даже для тех же значений PSNR).

На рис.6 представлено сравнение разработанного кодека со стандартным широко распространенным кодеком H.264 [10] и вейвлет-кодеком Dirac [11]. Сравнимые кодеки имеют следующие настройки: базовый профиль, только один ключевой кадр на всю последовательность, В-кадры и переменный размер блоков были отключены, использовалась дробнопиксельная компенсация по умолчанию, видеопоследовательности обрабатывались только в режиме оттенков серого.

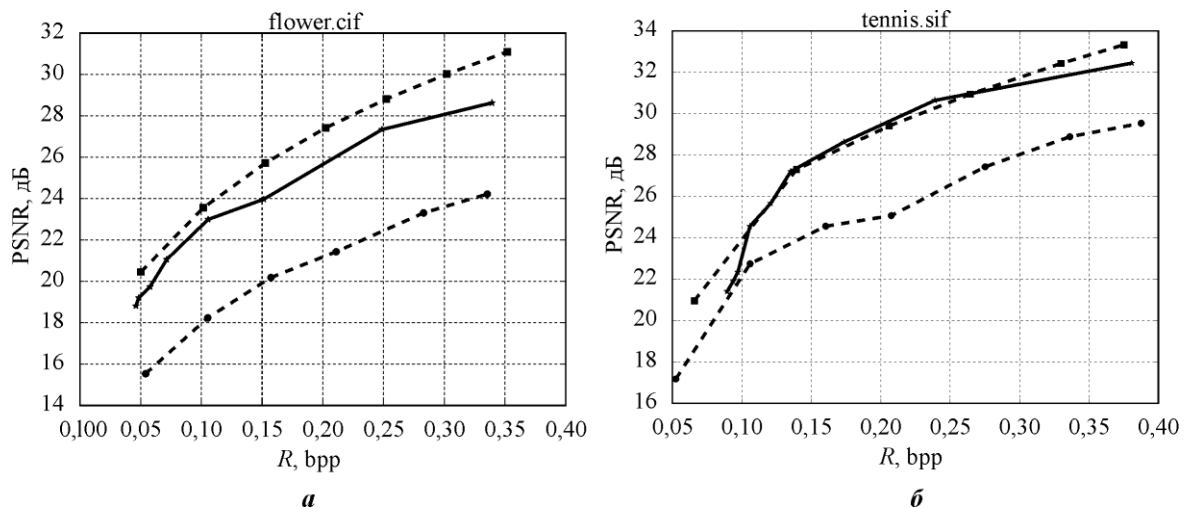


Рис.6. Сравнение кодеков с шестиугольной формой блоков (—), H.264 (■—■) и Dirac (●—●) на примере видеопоследовательностей Flower Garden (а) и Tennis (б)

Заключение. Результаты экспериментов подтверждают, что разработанный видеокодек на основе блоков шестиугольной формы практически всегда показывает лучшие результаты, чем видеокодек Dirac, и близок к результатам кодека H.264.

Литература

1. Ричардсон Я. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения. – М.: Техносфера, 2005. – 368 с.
2. Ватолин Д., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. – М.: ДИАЛОГ-МИФИ, 2002. – 384 с.
3. Умняшкин С.В., Шаронов И.О. Использование шестиугольных блоков для компенсации движения при видеокомпрессии // Цифровая обработка сигналов. – 2012. – №3. – С. 2–8.
4. Умняшкин С.В., Шаронов И.О. Способ блочной межкадровой компенсации движения для видеокодеков // Патент РФ № 2493670, приоритет от 15.12.2011.
5. Александров А.А., Умняшкин С.В. Модификация алгоритма трехслойной компенсации движения для видеокодека на базе дискретного вейвлет-преобразования // Изв. вузов. Электроника. – 2011. – №6. – С. 55–61.
6. Nelson M., Gailly J.-L. The data compression book. – 2nd Ed. – NY: M&T Books, 1996. – 576 p.
7. Said A., Pearlman W.A. A new fast and efficient image codec based on set partitioning in hierarchical trees // IEEE Transactions on Circuits and Systems for Video Technology. – June 1996. – Vol. 6. – P. 243–250.
8. Умняшкин С.В., Табориский А.Е. Метод обработки Р-кадров для видеокодека на основе дискретного вейвлет-преобразования // Цифровая обработка сигналов. – 2011. – №3. – С. 11–16.

9. *Antonini M., Barlaud M., Mathieu P., Daubechies I.* Image coding using wavelet transform // IEEE Transactions On Image Processing. – 1992. — Vol. 1. – № 2. – P. 205–220.

10. x264, сообщество VideoLAN. – URL: <http://www.videolan.org/developers/x264.html> (дата обращения: 01.07.2014).

11. Dirac Video Compression. – URL: <http://diracvideo.org> (дата обращения: 01.07.2014).

Статья поступила
10 марта 2015 г.

Умняшкин Сергей Владимирович – доктор физико-математических наук, профессор кафедры высшей математики №1 МИЭТ. *Область научных интересов:* цифровая обработка сигналов и изображений. **E-mail: vrinf@miec.ru**

Шаронов Игорь Олегович – аспирант кафедры высшей математики №1 МИЭТ. *Область научных интересов:* цифровая обработка сигналов и видеосжатие.

Информация для читателей журнала

«Известия высших учебных заведений. Электроника»

Вы можете приобрести в редакции журналы за 2014 г. и за первое полугодие 2015 г.

Адрес редакции: 124498, г. Москва, г. Зеленоград, пл. Шокина, д.1, МИЭТ, комн. 7231.

Тел.: 8-499-734-62-05. E-mail: magazine@miec.ru

<http://www.miet.ru/structure/s/894/e/12152/191>