

Алгоритм видеокомпрессии на основе дискретного вейвлет-преобразования с трехслойной схемой кодирования векторов движения

А.А.Александров, Е.А.Коплович, С.В.Умняшкин

Московский государственный институт электронной техники
(технический университет)

Предложен алгоритм видеокомпрессии, основанный на иерархическом кодировании коэффициентов дискретного вейвлет-преобразования и блочной компенсации движения с перекрытием. В рамках данного алгоритма реализована трехслойная схема кодирования векторов движения, которая позволяет формировать вложенный код и гибко регулировать распределение битовых затрат между разностным кадром и векторным полем.

В настоящее время в наиболее распространенных видеокодеках базовым инструментом, используемым для устранения внутрикадровой (пространственной) избыточности данных, является дискретное косинусное преобразование (ДКП). Это связано не только с большей сложностью реализации дискретного вейвлет-преобразования (ДВП) по сравнению с ДКП и не только с историческим аспектом (возможности применения ДКП для сжатия изображений начали изучаться примерно на 15 лет раньше ДВП). Основная проблема заключается в том, чтобы эффективно соединить в структуре алгоритма видеокомпрессии этапы устранения межкадровой (временной) и внутрикадровой избыточности данных.

Как правило, снижение временной избыточности достигается за счет предсказания нового кадра $\mathbf{I}$ на основе одного или нескольких ранее закодированных опорных кадров. Для построения прогнозного кадра $\mathbf{I}_{pred}$ чаще всего используется блочная компенсация движения в пространственной области как наиболее простая и эффективная. Если полученное разностное изображение

$$\mathbf{I}_{diff} = \mathbf{I} - \mathbf{I}_{pred} \quad (1)$$

обрабатывается с помощью блочного преобразования (в частности, ДКП), это приводит к тому, что при высоких уровнях сжатия на декодируемом изображении начинает проявляться соответствующая блочная структура, ухудшающая субъективную оценку его качества.

В отличие от ДКП, ДВП применяется ко всему кадру (ключевому или разностному) целиком, что позволяет избежать «разваливания» декодированного изображения на отдельные фрагменты. Однако использование ДВП для обработки разностного кадра (1), полученного после блочной компенсации движения, имеет существенный недостаток: такой кадр часто обладает ярко выраженной блочной структурой, что приводит к «всплескам» коэффициентов вейвлет-спектра, соответствующих границам блоков и, как следствие, к неэффективному сжатию.

Чтобы снизить влияние компенсации движения в пространственной области на качество компрессии, были предложены способы устранения межкадровой избыточности

с помощью вейвлет-коэффициентов, включая как самые простые алгоритмы, дающие сравнительно грубую оценку искомого вектора [1], но при этом имеющие низкие требования к вычислительной мощности и объему памяти, так и более сложные методы, для которых необходимы большие вычислительные ресурсы [2, 3]. Компенсация в области вейвлет-преобразования, как правило, дает лучшие результаты (в плане построения прогнозного кадра) по сравнению с обычной блочной компенсацией в пространственной области, но на текущий момент она не получила широкого распространения ввиду сложности реализации.

Компромиссом между качеством и эффективностью вычислений является использование *блочной компенсации с перекрытием* (overlapped block motion compensation – OBMC) [4]. Основная идея данного метода заключается в том, что блочные артефакты могут быть частично сглажены с помощью интерполяции прогнозных значений яркости кадра на основе векторов движения соседних блоков. Прогноз для блока из i -го кадра строится следующим образом:

$$\mathbf{I}_{pred(OBMC)}^i(y, x) = \sum_{k=1}^4 \mathbf{I}^{i-1}(y + y_{mv_k}, x + x_{mv_k}) \mathbf{W}(n, m),$$

где $y_{mv_k}, x_{mv_k}, k = 1, \dots, 4$ – найденные координаты векторов движения соседних блоков, $\mathbf{W}(n, m) = W(n)W(m)$ – весовая функция «окна». Идея метода проиллюстрирована на рис.1. Окно, наложенное на текущий макроблок размером $N \times N$, перекрывает восемь соседних блоков (рис.2). В результате прогноз для значения яркости пиксела из текущего блока зависит от четырех ранее найденных векторов движения: вектора для данного блока и трех векторов для соседних блоков.

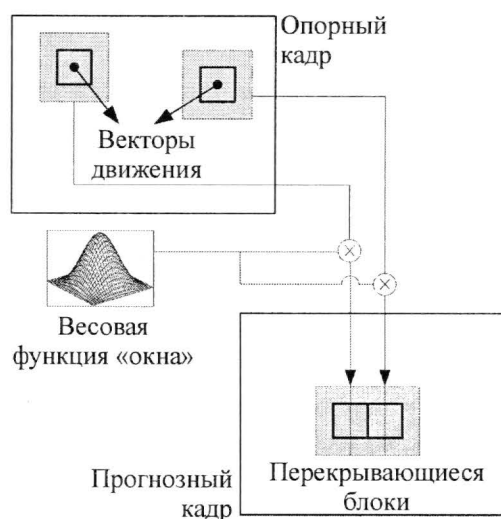


Рис.1. Построение прогноза с использованием перекрывающихся блоков

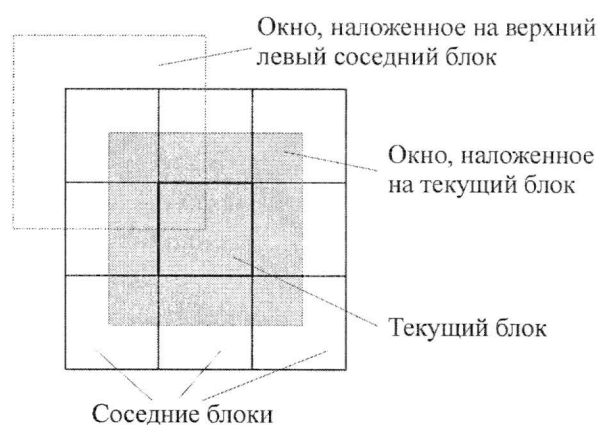


Рис.2. Пример перекрывающихся блоков

Для сжатия разностного изображения (1), построенного с помощью прогнозного кадра $\mathbf{I}_{pred(OBMC)}^i$, в предлагаемом видеокодере используется алгоритм SPIHT [5]. Метод SPIHT ориентирован на скорейшую передачу наиболее важной информации об изображении и позволяет генерировать полностью вложенный битовый код. Благодаря данному свойству видеокодер на основе алгоритма SPIHT может точно регулировать би-

товые затраты на кодирование разностного кадра, останавливая передачу данных, когда превышен выделенный битовый бюджет или лимит времени на кодирование кадра.

Если разностное изображение, как правило, сжимается с потерями, то информация о найденных векторах движения, необходимых для построения прогнозного кадра, должна передаваться без каких-либо изменений. Таким образом, встает вопрос о нахождении метода масштабируемого кодирования векторов движения. Особенно актуальной эта проблема становится при низких битовых затратах, так как в этом случае закодированные данные о движении могут занимать значительную часть ограниченного канала передачи, не оставляя битового бюджета для кодирования разностного кадра.

В предлагаемом алгоритме видеокомпрессии реализован метод трехслойного кодирования векторов движения. Первый слой является базовым, остальные два слоя – уточняющими. Количество используемых слоев может выбираться в зависимости от пропускной способности канала связи и доступных вычислительных ресурсов.

На рис.3 показано распределение блоков изображения по слоям. На первом слое определяются векторы перемещения только для четверти блоков, еще четверть блоков обрабатывается на втором слое, для оставшейся половины блоков векторы движения находятся на третьем слое.

Каждому блоку на текущем слое соответствует вектор перемещения, найденный с помощью RD-оптимизации, т.е. минимизации функции Лагранжа:

$$J = D_{MV} + \lambda R_{MV}, \quad (2)$$

где D_{MV} – искажение, вносимое ошибкой

предсказания блока; R_{MV} – количество бит, необходимое для кодирования вектора, λ – множитель Лагранжа, задающий баланс между качеством сжатия и битовыми затратами.

На каждом слое вектор перемещения кодируется как разность между найденным и прогнозным вектором. Прогнозный вектор определяется по ранее вычисленным векторам движения соседних блоков. Пусть $\mathbf{m}_A$, $\mathbf{m}_B$, $\mathbf{m}_C$, $\mathbf{m}_D$ – векторы перемещения соседних блоков, тогда прогнозный вектор находится с помощью покоординатной медианной фильтрации как

$$\mathbf{m}_{pred} = \text{median}(\mathbf{m}_A, \mathbf{m}_B, \mathbf{m}_C, \mathbf{m}_D). \quad (3)$$

На первом слое для прогноза используются векторы блоков, показанных на рис.4,а. Прогноз на втором слое строится с помощью соседних блоков из первого слоя (рис.4,б). Вектор для блока на третьем слое прогнозируется по двум векторам из первого слоя и двум векторам из второго слоя (рис.4,в).

На каждом слое разности между прогнозными и вычисленными векторами движения кодируются с помощью интервального кодера [6] с квазистатической моделью распределения. Накопленная по обработанным данным гистограмма частот появления различных векторов движения, используемая для статистического кодирования, «сбрасывается» в некоторое начальное состояние после кодирования ключевого кадра. В случае, когда не для всех блоков были определены векторы перемещения, т.е. ис-

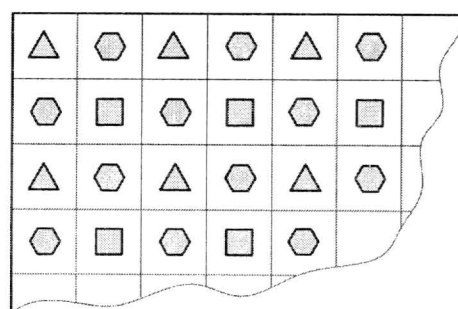


Рис.3. Распределение блоков изображения по слоям: $\square$ – первый слой; $\triangle$ – второй слой; $\hexagon$ – третий слой

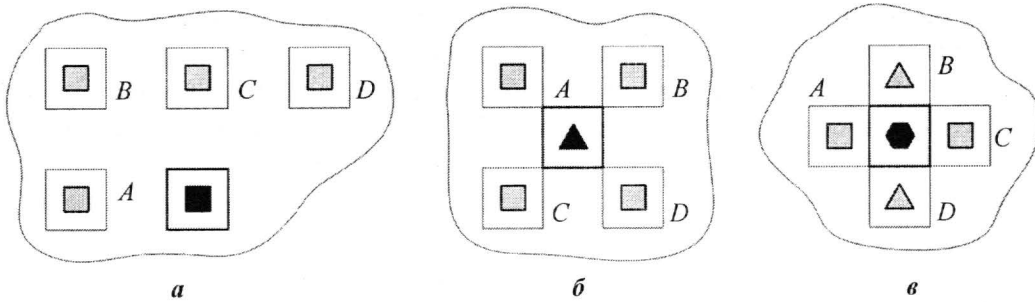


Рис.4. Блоки, используемые для предсказания вектора перемещения на первом (а), втором (б) и третьем (в) слоях

пользовалось менее трех слоев, недостающие векторы находятся с помощью медианной фильтрации (3) по ближайшим вычисленным соседям. Эти векторы не кодируются, так как декодер может восстановить их, выполнив аналогичную процедуру фильтрации. Если после кодирования очередного слоя битовый бюджет на кодирование информации о движении не был превышен, обрабатывается следующий слой.

Как было сказано ранее, векторы движения используются для построения прогнозного кадра, который затем вычитается из кодируемого кадра. Полученное разностное изображение переводится в область ДВП и кодируется алгоритмом SPIHT, при этом

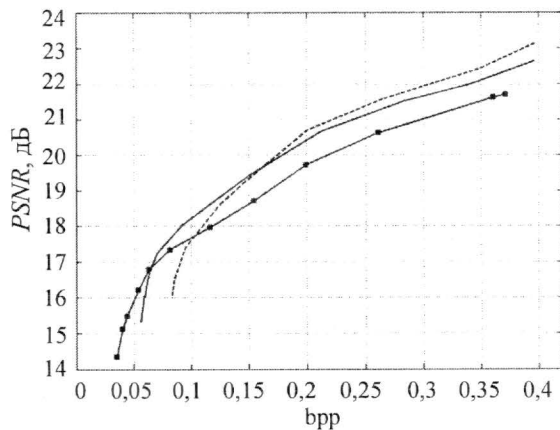


Рис.5. Результат кодирования тестовой последовательности Flower Garden: —•— один слой; — два слоя; - - - - три слоя

после кодирования каждой битовой плоскости для нахождения оптимального баланса между качеством и степенью сжатия вычисляется значение функции Лагранжа, аналогичной (2), с тем же параметром λ .

На рис.5 показана зависимость качества восстановленного видеоизображения от битовых затрат для стандартной тестовой видеопоследовательности Flower Garden. Данная последовательность длительностью 5 с состоит из 150 кадров размером 352×224 пиксела, первый кадр кодировался как ключевой, остальные — как разностные. Качество восстановленных кадров оценивалось с помощью пикового отношения сигнал-шум:

$$PSNR = 10 \lg \left(\frac{255^2}{\frac{1}{w \cdot h} \sum_{i,j} (I(i,j) - I'(i,j))^2} \right),$$

где w, h — размеры изображения; $I(i, j)$ и $I'(i, j)$ — яркости точек исходного и восстановленного кадров.

Из графика видно, что на низких битрейтах оправдано использование только одного слоя. Далее при увеличении доступного битового бюджета можно переходить к обработке следующих слоев.

Таким образом, предложенная трехслойная схема обработки блоков совместно с алгоритмом SPIHT в рамках вейвлет-видеокодера дает возможность более гибко распределять битовые затраты между разностным изображением и полем векторов движения. Повышение качества сжатия достигается за счет использования RD-оптимизации как на этапе нахождения векторов перемещения, так и при кодировании разностного кадра. Трехслойная схема проста в реализации и может быть легко встроена в любой алгоритм видеокодирования с блочным предсказанием движения.

Работа выполнена при поддержке гранта Президента Российской Федерации (МД-6172.2008.9).

Литература

1. **Joo Sanghyun, Kikuchi H.** A new motion compensation on a wavelet transform domain // Image Processing, 2000. Proceedings. 2000 International Conference on (10–13 Sept. 2000). – 2000. – Vol. 3. – P. 130 – 133.
2. **Hyun-Wook Park, Hyung-Sun Kim.** Motion estimation using low-band-shift method for wavelet-based moving-picture coding. Image Processing // IEEE Transactions on. – 2000. – Vol. 9, Issue 4. – April. – P. 577 – 587.
3. **Tjoa M.Y., Imanishi Y., Yamane N., Morikawa Y.** A motion compensation method using least squares motion estimation filter in wavelet domain. Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia // Proc. of the 2003 Joint Conf. of the Fourth International Conference on. – (15–18 Dec. 2003). – Vol. 1. – P. 46 – 50.
4. **Orchard M.T., Sullivan G.J.** Overlapped block motion compensation: an estimation-theoretic approach Image Processing // IEEE Transactions on. – 1994. – Vol. 3, Issue 5, Sept. – P.693 – 699.
5. **Said, W.A. Pearlman.** A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees // IEEE Transactions on Circuits and Systems for Video Technology. – 1996. – Vol. 6. – June. – P. 243 – 250.
6. **Martin G.N.N.** Range encoding: an algorithm for removing redundancy from a digitised message // Presented to the Video & Data Recording Conference. – Southampton. – 1979. – July 24–27.

Статья поступила
25 августа 2008 г.

Александров Андрей Александрович – стажер-исследователь НИИ ВСиСУ МИЭТ. *Область научных интересов:* цифровая обработка сигналов, в том числе сжатие данных без потерь, сжатие изображений и видеопоследовательностей.

Коплович Евгения Александровна – аспирант кафедры ВМ-1 МИЭТ. *Область научных интересов:* цифровая обработка сигналов, сжатие изображений и видеоданных.

Умняшкин Сергей Владимирович – доктор физико-математических наук, профессор, проректор по информатизации и международной деятельности МИЭТ. *Область научных интересов:* цифровая обработка сигналов и изображений.