

# ОСОБЕННОСТИ МЕЖКАДРОВОГО КОДИРОВАНИЯ ВИДЕОИНФОРМАЦИИ ПО СТАНДАРТУ MPEG-4

Власюк И.В., Сидорова А.И., Романова Е.П.

Аннотация. В докладе рассматривается стандарт сжатия видеoinформации MPEG-4. В работе приведена возможная реализация кодера, соответствующая рассматриваемому стандарту. Рассмотрена работа алгоритма компенсации движения, отмечены общие недостатки используемых алгоритмов формирования вектора движения макроблоков и предложен способ их устранения.

Появление DVD-проигрывателей, цифрового спутникового телевидения, новых цифровых камер и «веб-камер» - все это было бы невозможно без развития алгоритмов сжатия и обработки видео. Цель, которую преследуют стандарты сжатия - повышение степени сжатия видео, или повышение качества при той же степени сжатия. Например, стандарт MPEG-4 AVC/H.264 был принят в конце 2003 года и создавался с целью удвоить степень сжатия видео по сравнению с MPEG-4 (1998 г.).

Рассмотрим алгоритм кодирования видеоданных по стандарту MPEG-4. Обратимся к блок-схеме видеокодера (рис.1). В схеме можно выделить два основных пути прохождения сигнала: кодирование и реконструкция.

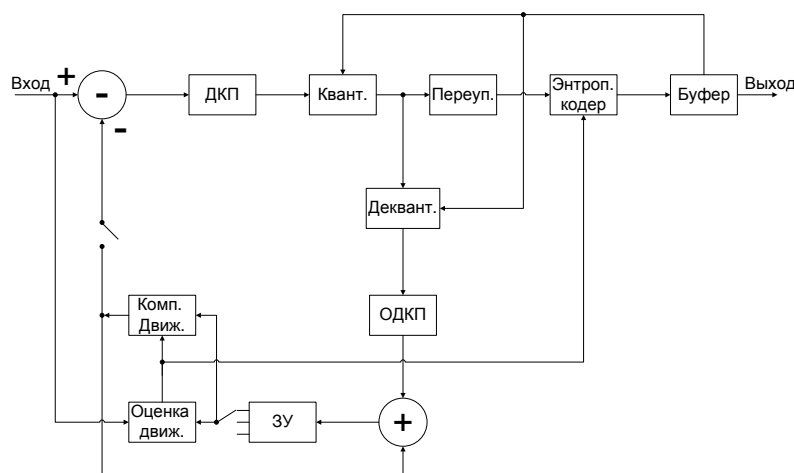


Рис.1. Блок-схема видеокодера по стандарту MPEG-4

Поток кодирования состоит в следующем:

На вход кодера поступает входной кадр  $F_n$ . Текущий входной кадр сравнивается со ссылочным кадром (блок оценки движения), в качестве которого может быть взят предыдущий закодированный кадр. Несколько предыдущих закодированных кадров хранятся в блоке зУ – запоминающем устройстве. Функция оценки движения находит на ссылочном кадре макроблок, наиболее похожий на выбранный макроблок текущего кадра, и смещение текущего макроблока относительно макроблока на ссылочном кадре фиксируется в виде вектора движения.

Затем функция компенсации движения строит кадр-прогноз  $P$  на основе найденного вектора движения и ссылочного кадра, взятого из ЗУ. Далее кадр-прогноз  $P$  вычитается из текущего кадра. Их разность  $D$  называется остаточным кадром.

Остаточный кадр поблочно преобразуется с помощью дискретного косинусного преобразования, далее каждый блок квантуется. Затем для коэффициентов преобразования блоков производится переупорядочение и результат, вместе с векторами движения, поступающими с выхода блока оценки движения, кодируется энтропийным кодером. Далее сигнал поступает на вход буферного запоминающего устройства, которое оценивает поток видеоданных. Кодек является адаптивным, т.к. буферное устройство регулирует поток данных: в зависимости от величины битового потока управляет количеством уровней квантования (блоки квантования и деквантования). На выходе буферного устройства получаем сжатый битовый поток.

Теперь вернемся к потоку реконструкции данных:

К каждому квантованному макроблоку применяется деквантование и обратное косинусное преобразование. Кадр-прогноз  $P$  (кадр с компенсированным движением) складывается с полученным остаточным кадром  $D'$ , в результате получаем реконструированный кадр  $F_n'$ . Полученный реконструированный кадр записывается в запоминающее устройство. Так как процесс квантования является необратимым, он вносит искажения, и реконструированный кадр  $F_n'$  не совпадает с текущим.

Реконструкция нужна для того, чтобы в кодере использовался точно такой же ссылочный кадр, как и в декодере, в этом случае в декодере не будет дрейфа ошибок от изображения к изображению.

Объем передаваемой информации сильно снижается, благодаря тому, что кодируется не каждый целый кадр видеопоследовательности, а разность между кадрами. Энергия кадра-остатка значительно меньше энергии кадра видеопоследовательности, что хорошо видно на рисунке 2.

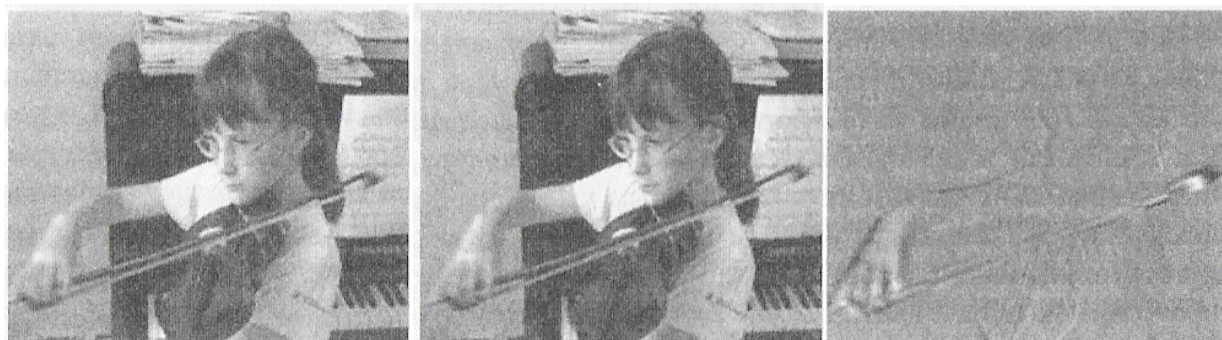


Рис.2. Кадры видеопоследовательности и их разность

На этом рисунке нейтральный серый цвет соответствует нулевым разностям пикселей, а более яркий или более темный оттенок серого цвета соответствует положительным или отрицательным разностям. Компенсация движения позволяют еще больше уменьшить энергию кадра-остатка.

Движущиеся объекты на видеосцене редко следуют строго вдоль границ блоков пикселей 16 x 16, поэтому более эффективным может быть использование переменных размеров блоков для оценки и компенсации движения.

В некоторых случаях лучший прогноз компенсации движения можно построить на основе интерполированных (промежуточных) положений сэмплов на ссылочном кадре. Подпиксельная оценка и компенсация движения использует при поиске промежуточные пиксельные позиции наряду с целыми позициями. Кодер ищет позицию, которая обеспечит наилучшее совпадение макроблоков (т.е. минимизирует энергию остатка) на основе целых и дробных значений позиций для прогноза компенсированного движения.

Рисунок 3 иллюстрирует оценку движения с четвертьпиксельной точностью. На первой стадии ищется лучшее совпадение текущего макроблока по целочисленной решетке сэмплов (круги). Затем кодер делает поиск по позициям полупикселей сразу после найденной позиции (квадраты) в целях улучшения совпадения, а затем, если необходимо, совершает поиск с шагом в четверть пиксела

(треугольники) для нахождения еще лучшего совпадения. Потом делается вычитание из текущего блока или макроблока его наилучшего прогноза, смещение (вектор компенсации движения) которого может быть кратно целым, половинам или четвертям пикселей.

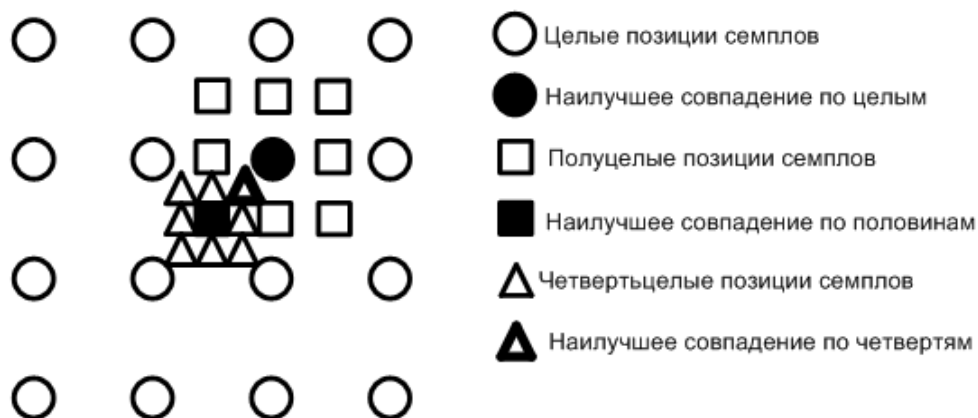


Рис.3. Четвертьпиксельная интерполяция

В общем случае более «тонкая» интерполяция дает лучшие характеристики компенсации движения (с меньшей разностью), но при существенном увеличении объема вычислений. Поиск подходящих блоков 4 x 4 с интерполяцией по четвертьпикселям является более сложной задачей по сравнению с использованием блоков 16 x 16 без интерполяции. Помимо существенного возрастания сложности вычислений приходится платить за

кодирование большего числа векторов движения для последующей реконструкции изображения. При уменьшении размера блока приходится передавать больше векторов. Также больше битов необходимо использовать для представления векторов движения по полупикселям и четвертьпикселям, так как приходится кодировать не только целые, но и дробные (т.е. 0,5 и 0,25) части компонент этих векторов.

Но даже в кодерах эффективных алгоритмов, таких как MPEG-4 и H.264, существует проблема неверного нахождения векторов движения. Пример таких ошибок можно увидеть на рисунке 4.



Рис.4. Пример ошибочно найденных векторов движения

На неподвижном фоне диктора тоже присутствуют вектора движения, хотя смещение макроблоков фона равно нулю.

Появление «не существующих» векторов не приводит к заметным глазу искажениям изображения, но информация о векторах является лишней и ненужной, а значит, ее наличие снижает эффективность сжатия видеоинформации кодером. Так как буферное устройство регулирует поток данных, управляя квантователем, то из-за излишней информации, количество уровней квантования может быть сокращено, что не лучшим образом скажется на качестве изображения.

На рисунке 5 показаны частотно-контрастные характеристики двух телевизионной камер (без коррекции).

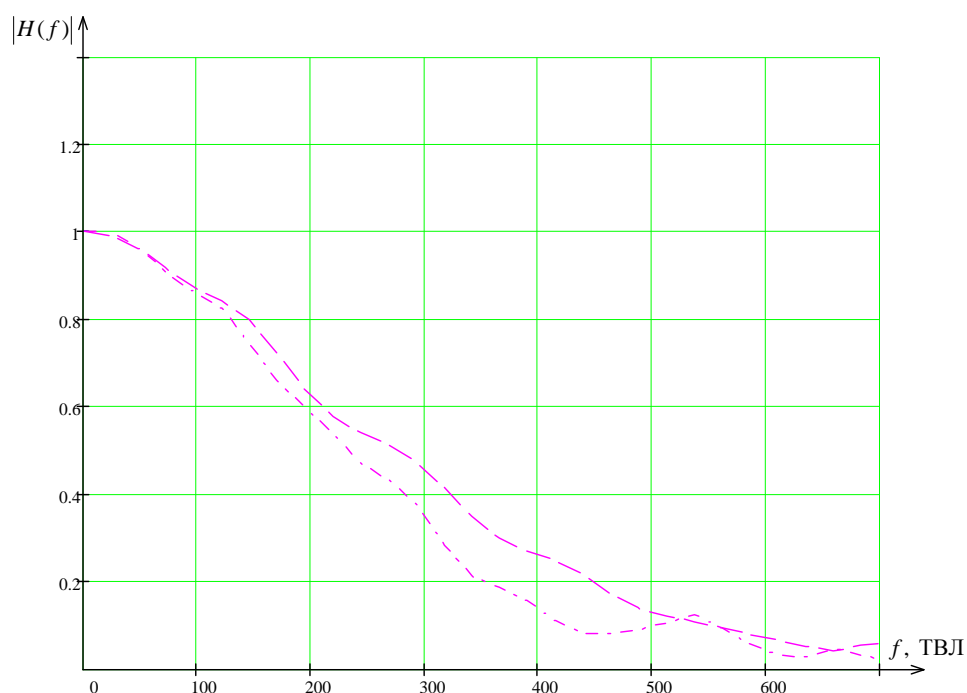


Рис.5. Частотно-контрастная характеристика (ЧКХ) ТВ камеры

По горизонтальной оси графика отложены пространственные частоты изображения. Данная характеристика показывает, что большая часть энергии ТВ сигнала сосредоточена в низкочастотной области. В области высоких частот наблюдается спад характеристики. Это говорит о том, что на высоких частотах изображения большее влияние имеют шумы, отношение сигнал/шум меньше, чем на низких частотах. Соответственно, чем мельче деталь изображения (чем больше пространственная частота), тем хуже коэффициент передачи.

Пространственные частоты изображения отражают коэффициенты косинусного преобразования, которое использует блоки, например, 8x8. Прямое преобразование FDCT (Forward DCT), примененное к блоку X размером NxN, задается следующей формулой:

$$Y = AXA^T \quad (1)$$

где X — это матрица сэмплов, Y — матрица коэффициентов, а A - NxN-матрица преобразования. Элементы матрицы A можно найти по формуле:

$$A_{ij} = C_i C_j \cos \frac{(2j+1)i\pi}{2N} \quad (2)$$

где  $C_i = \sqrt{\frac{1}{N}}$  ( $i = 0$ ),  $C_i = \sqrt{\frac{2}{N}}$  ( $i > 0$ ).

Выражения для X и Y можно переписать в виде сумм (3) и (4):

$$Y_{xy} = C_x C_y \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} X_{ij} \cos \frac{(2j+1)y\pi}{2N} \cos \frac{(2i+1)x\pi}{2N} \quad (3)$$

$$X_{ij} = C_x C_y \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} C_x C_y X_{xy} \cos \frac{(2j+1)y\pi}{2N} \cos \frac{(2i+1)x\pi}{2N} \quad (4)$$

Выходом двумерного FDCT служит матрица  $8 \times 8$  коэффициентов, которая представляет блок сэмплов изображения в области преобразования DCT. Как видно из рисунка 6, ненулевые коэффициенты DCT обычно образуют кластер вокруг левого верхнего коэффициента с примерно симметричным распределением вдоль диагонали блока. Левый верхний коэффициент – отражает постоянную составляющую (DC), правый нижний – самую высокую пространственную частоту.

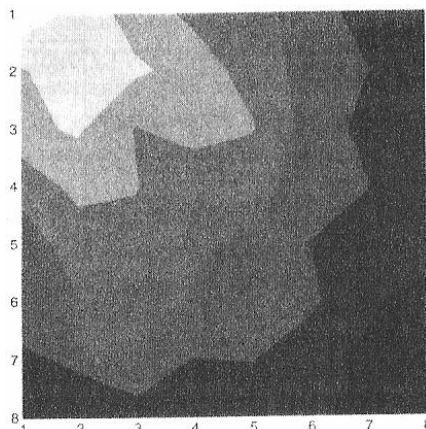


Рис.6. Коэффициенты ДКП

Таким образом, во избежание появления ошибок при нахождении векторов движения предлагается искать вектора иерархически. Так как хорошей точности нахождения векторов движения удастся достичь только при больших отношениях сигнал/шум, то сначала следует искать их в низкочастотной области. В этом случае, значительно снизится вероятность ошибки нахождения направления вектора, но точность нахождения местоположения блока будет небольшой. Положение блока возможно уточнить, используя высокочастотные составляющие. Из-за малой величины соотношения сигнал/шум на ВЧ ошибки нахождения величины вектора движения все равно возможны, но в данном случае отклонение будет не более чем на 8 элементов изображения. Такие ошибки незначительны и не приводят к весомой межкадровой разности, которая может снизить эффективность сжатия. Данный способ оценки движения может быть реализован и для новой версии стандарта MPEG-4 AVC(H.264).

#### Литература

1. Ватолин Д., Ратушняк А., «Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео». – М.: ДИАЛОГ-МИФИ, 2003, 384 с.;
2. Смирнов А. В., Пескин А. Е. «Цифровое телевидение: от теории к практике». – М.: Горячая линия – Телеком, 2005, 352 с.;
3. Ян Ричардсон, перевод с англ. Чепыжова В.В. «Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения», Москва: Техносфера, 2005, 368 с.;
4. И.В. Власюк., «Метод контроля пространственных характеристик телевизионных камер». Метрология и измерительная техника в связи. – 2005, – № 6. – С.13 – 16.