

Реализация IP-блока оценки векторов движения для кодека H.264 телевизионного сигнала высокой четкости

А.А. Беляев, И.Ю. Колесникова, Д.А. Кузнецов

ОАО НПЦ «ЭЛВИС», bel@elvees.com

Аннотация — Рассматриваются способы построения аппаратного ускорителя для решения задачи оценки векторов движения, являющейся одной из наиболее сложных с вычислительной точки зрения, при реализации сжатия телевизионного сигнала высокой четкости по стандарту H.264.

Ключевые слова — сжатие видеоданных, оценка вектора движения, стандарт H.264.

I. ВВЕДЕНИЕ

Для реализации современных стандартов сжатия телевизионного сигнала высокой четкости требуется колоссальная, пока еще недостижимая для процессоров общего назначения, производительность. В самом деле, один кадр формата Full HD (1080p) содержит примерно 2 миллиона (1920*1080) пикселей. При скорости 60 кадров в секунду это соответствует потоку 120 мегапикселей в секунду. Учитывая, что каждый пиксель состоит из трех цветовых компонент, полный поток видеоданных составляет 360 Мбайт/с. И это только входной поток! А ведь для реализации указанных алгоритмов требуется выполнять сотни и даже тысячи операций над каждым пикселем. По некоторым данным [1] для реализации сжатия видеоданных по стандарту H.264 [2, 3] требуется около 3 тысяч операций на пиксель. В рамках традиционных процессорных архитектур, ориентированных на последовательную обработку данных, это в настоящее время практически невыполнимо.

По этой причине в последние годы ведущие мировые фирмы-разработчики IP-блоков, такие как Imagination, CEVA и ряд других, наращивают свои усилия в направлении разработки специализированных аппаратных ускорителей, реализующих наиболее популярные стандарты сжатия видеоданных – MPEG2, MPEG4, H.264 [4]. В состав всех этих ускорителей входит блок оценки вектора движения, на долю которого приходится, как правило, не менее 50% всей вычислительной нагрузки при реализации указанных методов сжатия. Таким образом, блок оценки вектора движения является ключевой составной частью любого из перечисленных видеокодеков.

В настоящей статье дается сравнительный анализ различных алгоритмов работы и способов построения блока оценки векторов движения, предназначенного

для реализации сжатия телевизионного сигнала высокой четкости (Full HD) в соответствии с одним из наиболее сложных, с вычислительной точки зрения, современных стандартов сжатия видеоданных – стандартом H.264.

II. ПРИНЦИПЫ РАБОТЫ БЛОКА ОЦЕНКИ ВЕКТОРА ДВИЖЕНИЯ

Базовой единицей кодирования в стандарте H.264 является макроблок – область изображения размером 16x16 пикселей. Для представления исходного изображения в стандарте H.264 используется цветовое пространство YCbCr с субдискретизацией (либо без субдискретизации) цветовых компонент – форматы 4:2:0, 4:2:2, 4:4:4.

В большинстве аппаратных и программных кодеров для оценки векторов движения используется только яркостная компонента изображения Y.

Вектор движения определяется как вектор смещения между положением некоторой области изображения (в данном случае макроблока или его части) на двух следующих друг за другом кадрах видеопоследовательности.

Вектора движения в стандарте H.264 имеют разрешение в четверть пикселя для компоненты яркости и в одну восьмую пикселя для компонент хроматичности. Таким образом, стандарт H.264 предъявляет повышенные требования к точности определения векторов движения по сравнению с предшествующими стандартами MPEG-2 (в котором вектора движения определяются с точностью до целых пикселей) и MPEG-4 (в котором используется полупиксельная точность).

Примеры определения вектора движения показаны на рис. 1. На приведенном рисунке блок размером 4x4 текущего кадра (а) прогнозируется по области ссылочного кадра, имеющего близкую позицию к положению текущего блока. Если горизонтальные и вертикальные компоненты вектора движения являются целыми числами (б), то соответствующий прогнозный блок пикселей существует в явном виде на ссылочном кадре (серые кружки). Если один или оба компонента вектора движения являются дробными числами (в), то блок-прогноз (серые кружки) генерируется с помощью

интерполяции по соседним пикселям на ссылочном кадре (белые кружки).

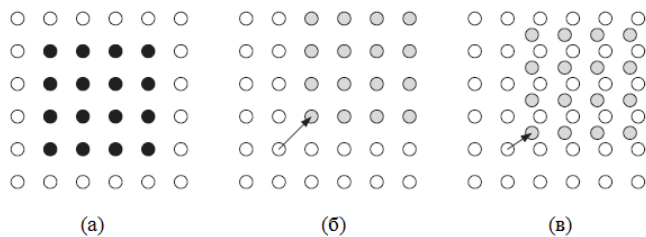


Рис. 1. Примеры прогноза по целым и дробным пикселям

В случае дробных векторов движения вычисление предсказанных значений компонент яркости и цветности производится с помощью интерполяции соседних с ними пикселей. Алгоритм интерполяции яркостной компоненты по полупикселям проиллюстрирован на рис. 2.

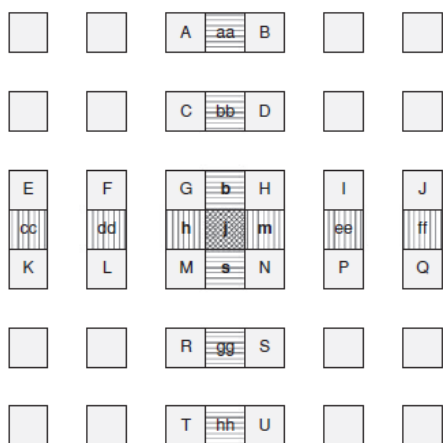


Рис. 2. Интерполяция компоненты Y по полупикселям

Полупиксельная интерполяция производится по шести соседним точкам, сначала по одной координате, затем по другой, согласно формулам:

$$b = \text{round}((E - 5F + 20G + 20H - 5I + J) / 32);$$

$$h = \text{round}((A - 5C + 20G + 20M - 5R + T) / 32);$$

$$j = \text{round}((cc - 5dd + 20h + 20m - 5ee + ff) / 32);$$

$$m = \text{round}((B - 5D + 20H + 20N - 5S + U) / 32);$$

$$s = \text{round}((K - 5L + 20M + 20N - 5P + Q) / 32).$$

Интерполированные значения четвертьпикселей вычисляются как полусумма прилегающих полупикселей или пикселей с целочисленными координатами, как это показано на рис. 3.

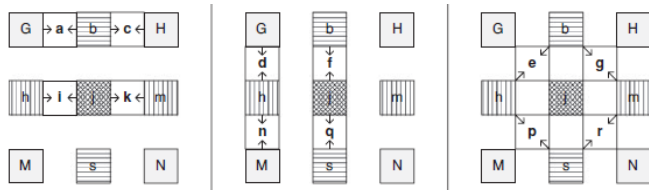
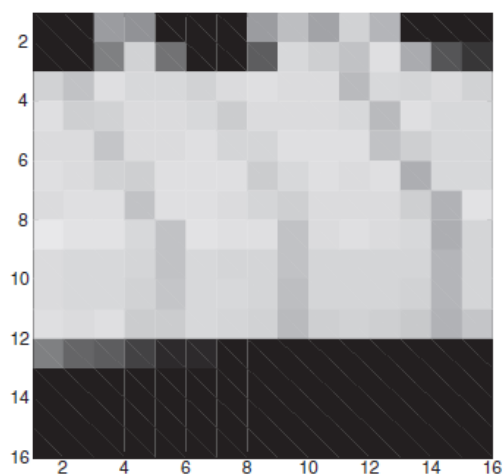
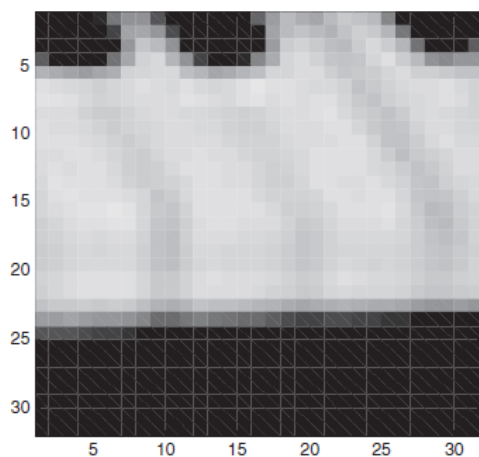


Рис. 3. Интерполяция компоненты Y по 1/4-пикселям

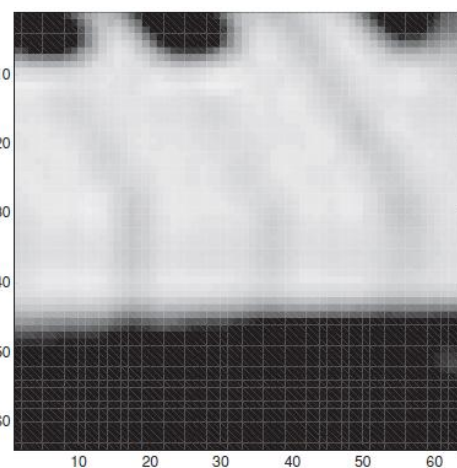
На рис. 4 показан результат интерполяции фрагмента изображения размером 16x16 пикселей (а), с 1/2-пиксельной (б) и 1/4-пиксельной (в) точностью.



а)



б)



в)

Рис. 4. Результаты интерполяции изображения: (а) исходный фрагмент; (б) результаты интерполяции с 1/2-пиксельной точностью; (в) результаты интерполяции с 1/4-пиксельной точностью

III. ВЫБОР АЛГОРИТМА ОПРЕДЕЛЕНИЯ ВЕКТОРА ДВИЖЕНИЯ

Алгоритмы определения вектора движения можно разделить на две неравные группы: к первой относятся алгоритмы полного перебора всех возможных значений вектора движения в пределах ссылочной области; вторая, более многочисленная, включает в себя различные алгоритмы быстрого поиска.

Алгоритм полного перебора позволяет гарантированно минимизировать ошибку предсказания, однако его применение связано с очень значительными (зависящими от размеров ссылочной области) вычислительными затратами.

Алгоритмы быстрого поиска дают возможность сократить объем вычислений, правда, за счет некоторого снижения качества предсказания. В табл. 1 приведены основные характеристики - пиковое отношение сигнал/шум (PSNR) и среднее количество точек поиска на макроблок (Points/MB) для некоторых известных алгоритмов [5]: алгоритм полного поиска (FS, full search), алгоритм «поиска бриллиантов» (DS, diamond search), улучшенный алгоритм «поиска бриллиантов» (IDS, improved diamond search), алгоритм 3-шагового поиска (TSS, three-step search), алгоритм 4-шагового поиска (FSS, four-step search), новый алгоритм 3-шагового поиска (NTSS, three-step search), 2-мерный алгоритм логарифмического поиска (2-DLS, 2-D logarithmic search).

Таблица 1

Пиковое отношение сигнал/шум (PSNR) и среднее количество точек поиска на макроблок (Points/MB) для различных алгоритмов оценки

| Алгоритм | FS | DS | IDS | TSS | FSS | NTSS | 2-DLS |
|-----------|-------|------|------|------|------|------|-------|
| PSNR, дБ | 29.5 | 29.4 | 28.9 | 27.0 | 29.3 | 28.7 | 26.8 |
| Points/MB | 207.4 | 18.6 | 9.7 | 23.7 | 19.7 | 21.6 | 16.0 |

Однако при аппаратной реализации очень важное значение имеет регулярность алгоритма поиска и возможность распараллеливания вычислений. С этой точки зрения алгоритм полного поиска оказывается более предпочтительным, что и предопределило в данном случае выбор в его пользу.

В рассматриваемом в данной работе устройстве применен модифицированный вариант алгоритма полного поиска – алгоритм «быстрого полного поиска» (FFS, fast full search), в котором поиск вектора движения выполняется в два этапа.

Обход по целочисленным координатам производится по спирали, как это показано на рис. 5. Началом обхода является точка с координатами (0,0), затем – (0,1), (1,1), (1,0), (1,-1), (-1, -1), (-1, 0), (-1,0) и так далее. На втором этапе алгоритма производится уточнение найденных на первом этапе целочисленных векторов движения в их окрестностях с дробными координатами, при этом найденное на первом этапе разбиение не изменяется.

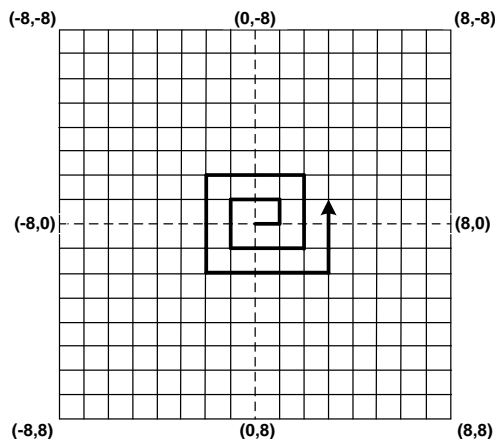


Рис. 5. Порядок обхода поля ссылочного фрагмента по целочисленным координатам

Порядок обхода окрестностей найденных на первом этапе целочисленных векторов движения показан на рис. 6. Началом обхода является точка с координатами (0,0), затем – (0,1/4), (1/4,1/4), (1/4,0), (1/4,-1/4), (-1/4, 0), (-1/4,-1/4), (-1/4,0) и так далее.

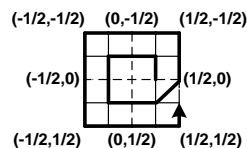


Рис. 6. Порядок обхода поля ссылочного фрагмента по дробным координатам

IV. ОСОБЕННОСТИ РАЗБИЕНИЯ МАКРОБЛОКОВ НА БЛОКИ В СТАНДАРТЕ H.264

В отличие от ранее существовавших стандартов H.263, MPEG-2, MPEG-4, в стандарте H.264 имеется возможность независимого определения вектора движения для блоков размером от 4x4 до 16x16, как это показано на рис. 7.

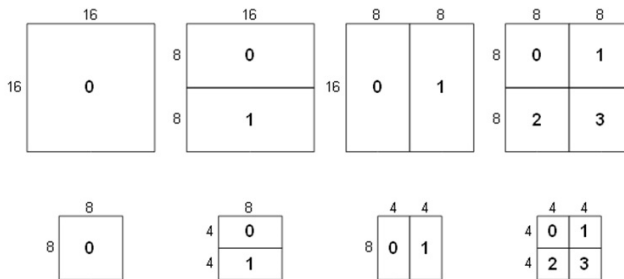


Рис. 7. Разбиения макроблока на блоки в стандарте H.264

Возможны разбиения макроблока на блоки с размерами: 16x16, 16x8, 8x16, 8x8, 8x4, 4x8 и 4x4. Всего существует 41 вариант (тип) таких блоков. Это значительно усложняет задачу, поскольку, в общем случае, вектор движения следует определить вначале для каждого из 41 типа блоков, после чего выбрать наилучшее разбиение макроблока на блоки, минимизирующее ошибку предсказания. Это

существенно замедляет работу «быстрых» алгоритмов, напротив, применение полного перебора с учетом этого становится еще более оправданным.

V. АППАРАТНАЯ СТРУКТУРА IP-БЛОКА ОЦЕНКИ ВЕКТОРОВ ДВИЖЕНИЯ И РЕЗУЛЬТАТЫ ЕГО ФИЗИЧЕСКОЙ РЕАЛИЗАЦИИ

Аппаратная структура разработанного IP-блока приведена на рис. 8.

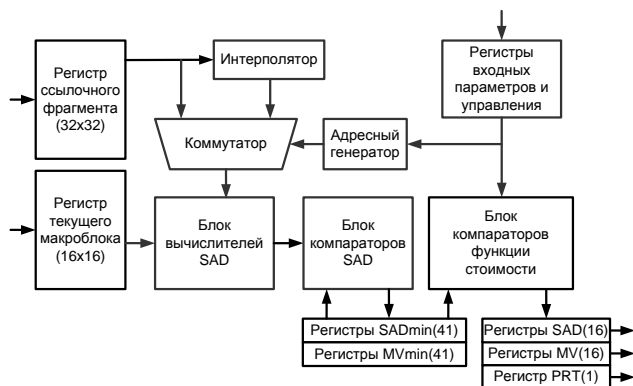


Рис. 8. Структурная схема IP-блока оценки векторов движения

В состав IP-блока входят: регистр ссылочного фрагмента; регистр текущего макроблока; интерполятор; коммутатор; блок вычислителей сумм поэлементных абсолютных разностей (SAD); адресный генератор; блок компараторов SAD; блок компараторов функций стоимости; регистры входных параметров и управления; набор регистров для хранения сумм поэлементных абсолютных разностей, векторов движения и разбиений макроблока.

IP-блок реализует алгоритм полного поиска. Ссылочная область изображения, хранящаяся в соответствующем регистре, имеет размер 32x32 пикселя. В зависимости от текущего смещения, значение которого определяется адресным генератором, из ссылочной области при помощи широкого 256-канального коммутатора выбирается фрагмент изображения размером 16x16 пикселей для сравнения с текущим макроблоком.

На первой стадии, при определении целочисленных значений координат вектора движения, предсказываемый фрагмент выбирается непосредственно из ссылочной области; на второй стадии, при определении дробных значений вектора, соответствующий фрагмент поступает с выхода интерполятора. После этого вычисляется сумма поэлементных абсолютных разностей между текущим макроблоком и предсказываемым фрагментом для всех 41 типов блоков, входящих в состав макроблока.

Полученное значение SAD для каждого типа блока сравнивается с ранее найденным минимальным значением, и если полученное значение оказывается меньше, то оно становится новым минимальным

значением, и так продолжается до окончания обхода всей ссылочной области. По окончании обхода, с использованием функции стоимости, производится окончательное определение разбиения макроблока на блоки и фиксация соответствующих значений векторов движения.

Предложенная схема IP-блока была физически реализована по технологии 40 нм (TSMC 40LP) в составе спроектированной в ОАО НПЦ «ЭЛВИС» многоядерной системы на кристалле MCom-02. При тактовой частоте 400 МГц IP-блок выполняет сжатие по стандарту H.264 видеоданных в формате Full HD со скоростью 120 кадров в секунду. Физические характеристики IP-блока приведены в табл.2.

Таблица 2

Физические характеристики IP-блока оценки векторов движения

| Характеристика | Значение |
|--|----------|
| Общая площадь, мкм ² | 540676 |
| Площадь комбинационных элементов, мкм ² | 448434 |
| Площадь некомбинационных элементов, мкм ² | 92242 |
| Общее количество регистров | 20783 |

VI. ВЫВОДЫ

На основе проведенного анализа предложена аппаратная структура IP-блока оценки векторов движения, предназначенного для применения в составе видеокodeка, выполняющего компрессию телевизионного сигнала высокой четкости в соответствии со стандартом H.264. Разработанный IP-блок производит оценку вектора перемещения макроблока размером 16x16 пикселей текущего кадра относительно ссылочного кадра с точностью до ¼ пикселя. Оценка производится независимо для всех входящих в макроблок блоков размером 16x8, 8x16, 8x8, 8x4, 4x8 или 4x4 пикселей. Разработанный IP-блок, физически реализованный по 40-нанометровой технологии (TSMC 40LP), при рабочей частоте 400 МГц позволяет выполнять сжатие по стандарту H.264 телевизионного сигнала высокой четкости (1080p) со скоростью 120 кадров в секунду.

ЛИТЕРАТУРА

- [1] Kempf T., Ascheid G., Leupers R. Multiprocessor Systems-on-Chips. Design Space Exploration. Springer, 2011.
- [2] ITU-T Recommendation H.264 and ISO/IEC 14496-10 AVC. 2011.
- [3] Richardson, Iain E. The H.264 advanced video compression standard. 2nd ed. John Wiley & Sons Ltd. 2010.
- [4] CEVA-MM3000. Product Brief. CEVA, Inc., 2011.
- [5] Liu Yiming, Li Haibiao, Xiao Yongjun. Improved Fast Motion Estimation Algorithm Based on the H.264 // International Forum on Information Technology and Applications. 2010.