

УДК 004.932

## РЕАЛИЗАЦИЯ МНОГОФУНКЦИОНАЛЬНОГО КОДЕРА AVC

*Гулевич А.Е., ФГУП «Главный радиочастотный центр», alex\_gulevich@mail.ru**Дворкович А.В., д.т.н., ФГУП «Главный радиочастотный центр», a\_dvork@nircom.ru**Кочарян А.Э., ФГУП «Главный радиочастотный центр», ashotkoch@mail.ru**Мингазов И.Д., ФГУП «Главный радиочастотный центр», elijacontact@gmail.com*

**Ключевые слова:** кодер, стандарт, изображение, компрессия, компенсация движения, пиксел, предсказание, обработка, фильтрация, количество бит, видеотрансляция.

Современный стандарт видеокодирования H.264/AVC был разработан Совместной Видео группой (Joint Video Team), образованной Группой Экспертов по Видеокодированию МСЭ-Т (VCEG ITU-T) и Экспертной Группой по Движущимся Изображениям МСС (MPEG ISO/IEC). Он обозначается как Рекомендация H.264 МСЭ-Т и Стандарт ISO/IEC 14496-10 (MPEG-4 Part 10 AVC) в МСС (последняя редакция стандарта – март 2009г. [1]). Основными целями разработки стандарта H.264/AVC были заявлены повышенная эффективность компрессии и обеспечение удобного для транспортировки по различным сетям представления видео как для интерактивных, так и вещательных приложений [2]. Стандарт H.264/AVC существенно повысил эффективность видеокодирования по отношению к ранее принятым стандартам MPEG-2 Part 2 [3], MPEG-4 Part 2 [4], H.263 [5].

Повышение эффективности видеокодирования в стандарте H.264/AVC обеспечивается применением как уже известных методов и подходов, так и новых функциональных возможностей: компенсация движения с использованием переменных размеров блока, включая малые размеры блока; компенсация движения с точностью до четверти пиксела; вектора движения, выходящие за границы изображения; компенсация движения с несколькими опорными изображениями; независимость порядка воспроизведения изображений и порядка опорных изображений; независимость методов обработки изображений и возможности их использования для предсказания движения; взвешенное предсказание; улучшенная обработка «пропущенных» (skipped) блоков и блоков с «прямым» (direct) предсказанием движения; направленное пространственное предсказание для внутрикадрового кодирования; деблокировочная фильтрация в цикле кодирования; преобразование блоков небольшого размера; иерархическое преобразование блоков; преобразование с использованием 16-битной арифметики; точное обратное преобразование; арифметическое или контекстно-адаптивное энтропийное кодирование. В стандарте H.264/AVC повышена устойчивость к ошибкам и потерям данных, обеспечивается гибкость работы на множестве сетевых структур.

*В статье рассмотрена реализация программного кодера стандарта H.264/AVC для систем вещания и двусторонней видеосвязи. Подробно рассмотрен алгоритм управления скоростью выходного потока кодера.*

Существенным шагом вперед стала реализация методов кодирования с масштабированием (Scalable Video Coding) в этом стандарте. По сравнению с предшествующими стандартами значительно повышена эффективность кодирования и снижены требования к производительности декодера.

Авторами был разработан программный видеокодер H.264/AVC. Этот кодер использован в системах VPhone (система видеоконференцсвязи и вещания), AVIS (система узкополосного наземного цифрового мобильного аудио- и видеовещания). При разработке данных систем к ним предъявлялись требования работы с каналами ограниченной пропускной способности, низкой задержки, значительного сжатия при обеспечении максимально возможного качества видеoinформации. Кодер данного стандарта был выбран, как способный удовлетворить заданным требованиям.

В разработанном авторами программном кодере H.264/AVC реализованы основные важные возможности стандарта, использование которых, с одной стороны, существенно повышает качество кодирования, а с другой стороны не приводит к резкому росту вычислительных затрат. Кодер не поддерживает взвешенное предсказание, избыточные слои, разделение данных, High профиль, SP/SI синхронизацию/переключение изображений, кодирование с масштабированием. Разработанный кодер может функционировать в составе ТВ кодера стандартного разрешения реального времени, в составе системы видеотрансляций в Интернете, в составе систем видеоконференцсвязи.

Важнейшей частью кодера при использовании в различных приложениях является блок контроля потока. При фиксированном параметре квантования количество бит каждого закодированного макроблока изменяется в зависимости от содержания кадра, поэтому скорость потока на выходе кодера существенно неравномерна. Обычно кодер с постоянными параметрами производит больше бит, если в видеосюжете присутствует быстрое движение и/или мелкие детали, и меньше бит - в случае медленного движения и/или отсутствия мелких деталей.

Вариации скорости потока могут вызывать проблемы для многих практических задач доставки и хранения



видео. Например, канал постоянной пропускной способности не может передавать поток переменной скорости, если флуктуации скорости потока превышают пиковую пропускную способность канала. В таких случаях происходит либо потеря данных, либо накопление задержки. Таким образом, необходимо адаптировать или контролировать скорость потока, производимого видеокодером, для удовлетворения требований, накладываемых пропускной способностью канала и механизмом доставки. Другое ограничение связано с фиксированным объемом информационных носителей, поэтому для оптимального заполнения объема носителя также необходимо контролировать скорость потока закодированного видео.

Неравномерность скорости потока, производимого видеокодером, может быть сглажена буферизацией закодированных данных перед передачей. Поток переменной скорости, создаваемый кодирователем, помещается в FIFO-буфер, этот буфер освобождается с постоянной скоростью, которая равна пропускной способности канала (данные из буфера помещаются в канал). Другой FIFO-буфер размещается перед входом декодера, он заполняется данными из канала с постоянной скоростью (равной пропускной способности канала) и освобождается декодером с неравномерной битовой скоростью, но с постоянной кадровой скоростью. Таким образом, поток переменной скорости может быть адаптирован к постоянной скорости канала посредством использования буферов кодирователя и декодера. Но такая адаптация достигается ценой внесения задержки и использования памяти для буферов, и чем больше вариации скорости потока, тем требуются большие задержка и размеры буферов. При произвольных вариациях скорости потока этот метод невозможно использовать на практике, т.к. это может привести к недопустимо высоким задержкам и размерам буферов, а система передачи зачастую должна иметь обратную связь.

Контроль потока управляет параметрами кодирователя (шагом квантования) с целью получения заданной скорости потока и минимизации искажений декодированного видео при данной скорости. Достижение оптимального компромисса между скоростью потока и качеством - нетривиальная задача, здесь могут применяться различные подходы и алгоритмы, в зависимости от типа видеоприложения. Их можно классифицировать следующим образом:

- Кодирование не в реальном времени на носители. Время кодирования жестко не ограничено, поэтому могут применяться сложные алгоритмы. Целью является помещение закодированного видео на доступный объем носителя, при этом максимизируется качество и, возможно, учитывается, что буфер декодирующего устройства или программы не должен переполняться или опустошаться в процессе декодирования. Могут применяться двухпроходные схемы (во время первого прохода кодирователь собирает статистику о видеопоследовательности, кодирование производится во время второго прохода).
- Кодирование живого видео для вещания. Процесс декодирования и буферизации может иметь ограни-

чения, в то время как кодирование может использовать мощное оборудование. Обычно допустима задержка в несколько секунд. Здесь применяются алгоритмы контроля потока средней сложности, возможно, включая двухпроходное кодирование каждого кадра.

- Кодирование для двусторонней видеоконференции. Каждый терминал производит как кодирование, так и декодирование, и вычислительная мощность может быть ограничена. Задержка должна быть минимальной (обычно в пределах 0,1 – 0,8 секунды). Здесь обычно применяются алгоритмы контроля потока невысокой сложности. Размеры буферов кодирователя и декодера должны быть минимизированы для обеспечения минимальной задержки, кодирователь должен точно выдерживать скорость выходного потока. Качество декодированного видео может значительно варьироваться, например, возможно существенное ухудшение качества при быстром движении, смене сюжета и т.п. Однако для некоторых задач (видеонаблюдение, телеконсультации) важно сохранять достаточное высокое качество изображения, при этом возможно пожертвовать постоянной кадровой скоростью. В этом случае применяется механизм пропуска кадров, он позволяет удовлетворить одновременно требования низкой задержки и постоянной, достаточно высокой четкости изображения.

В разработанном кодере H.264/AVC реализованы все перечисленные варианты регулирования выходного потока.

Алгоритм контроля потока разработанного кодирователя H.264/AVC построен на базе схемы SRC (Scalable Rate Control) [6-8]. Алгоритм выдерживает заданную скорость потока на сегменте кадров. Он использует следующую модель для скорости потока:

$$R = S \cdot \left( \frac{X_1}{Q} + \frac{X_2}{Q^2} \right), \quad (1)$$

где  $R$  – скорость потока,  $Q$  – шаг квантования,  $S$  – мера сложности кадра (в данной реализации – средняя абсолютная разность остаточного кадра после компенсации движения),  $X_1$  и  $X_2$  – параметры модели. Алгоритм состоит из следующих шагов, которые выполняются после компенсации движения для  $i$ -того кадра:

1. Вычисление целевой скорости потока  $R_i$ , основываясь на количестве кадров в сегменте кадров, числе бит, доступных для остатка сегмента кадров, и оценочной сложности  $i$ -го кадра (в качестве оценки используется размер предыдущего кадра). Если предыдущий кадр имел высокую сложность, то предполагается, что следующий кадр тоже будет сложным и должен занять большое число бит. Алгоритм пытается найти баланс между этим требованием и лимитом на число бит для сегмента:

$$R_{i+1} = \frac{G_i}{N_i} \cdot (1 - W) + A_i \cdot W, \quad (2)$$

где  $R_{i+1}$  – целевая скорость потока для  $i+1$  кадра,  $G_i$  – число бит, доступных для остатка сегмента кадров на момент времени  $i$ ,  $N_i$  – число кадров в остатке сегмента кадров на момент времени  $i$ ,  $A_i$  – число бит, использованных для кодирования  $i$ -того кадра,  $W$  – взвешивающий параметр.

Далее учитываются требования размера буфера:

$$\text{Если } B_{i+1} + R_{i+1} > 0.9 \cdot V, \text{ то } R_{i+1} = 0.9 \cdot V - B_{i+1}, \quad (3)$$

где  $B_i$  – заполненность буфера,  $V$  – размер буфера. Цель данного выражения – снизить вероятность переполнения буфера и выхода за пределы заданной задержки.

2. Вычисление шага квантования  $Q_i$  (вычисляется  $S$  для всего остаточного кадра и решается уравнение 1).

3. Кодирование кадра.

3. Обновление параметров модели  $X_1, X_2$ . Используется количество бит, фактически затраченных на кодирование  $i$ -того кадра  $R_i$ . Для этого выбираются точки данных, используя окно, размер которого зависит от изменения сложности видео. Если сложность изменяется значительно, то используется небольшое окно с предыдущими данными. При небольшом изменении сложности используется больше предыдущих данных. Далее методом линейной регрессии вычисляются параметры модели  $X_1, X_2$ :

$$X_2 = \frac{n \sum_{i=1}^n R_i - (\sum_{i=1}^n Q_i^{-1})(\sum_{i=1}^n Q_i R_i)}{n \sum_{i=1}^n Q_i^{-2} - (\sum_{i=1}^n Q_i^{-1})^2},$$

$$X_1 = \frac{(\sum_{i=1}^n Q_i R_i) - X_2 \sum_{i=1}^n Q_i^{-1}}{n} \quad (4)$$

После вычисления новых параметров из рассмотрения выбрасываются статистически «плохие» точки, для которых разность теоретически и практически затраченных бит превосходит величину стандартного отклонения. Таким образом, получается новый, более репре-

зентативный набор данных, для которого еще раз пересчитываются параметры модели по формуле 4.

Результаты работы приведенного алгоритма проиллюстрированы на рис. 1. Была использована последовательность кадров из фильма «Матрица» размера CIF, 25 кадров/сек, задержка кодирования не более 0,25 сек, выходной поток 300 кбит/сек. По горизонтали отложены номера кадров, размеры кадров – в байтах, задержка – в секундах. Из рис. 1 видно, что в районе 370-го кадра (смена сюжета) задержка резко выросла (более 3 сек), то есть механизм управления скоростью потока обеспечивает нужный поток, но не обеспечивает задержку.

Для предотвращения переполнения буфера изменяется механизм пропуска кадров: если кодер предсказывает, что следующий кадр вызовет переполнение буфера, то следующий кадр пропускается. Это эффективный метод защиты от переполнения буфера, но он может вызвать значительное снижения качества восприятия видео в случае пропуска нескольких последовательных кадров. Для борьбы с этой проблемой алгоритм пропуска кадров несколько модифицируется и выглядит следующим образом. Перед кодированием следующего кадра, кодер вычисляет заполненность буфера и оценивает размер следующего кадра (используется размер предыдущего). Если их сумма выше некоторого порога (например, 80% от размера буфера), то следующий кадр пропускается. Данный порог применяется для снижения вероятности пропуска группы последовательных кадров. Он может быть адаптивным или являться некоторой предопределенной константой.

На рис. 2 проиллюстрирована работа такого модифицированного алгоритма. Теперь заданная задержка выдерживается.

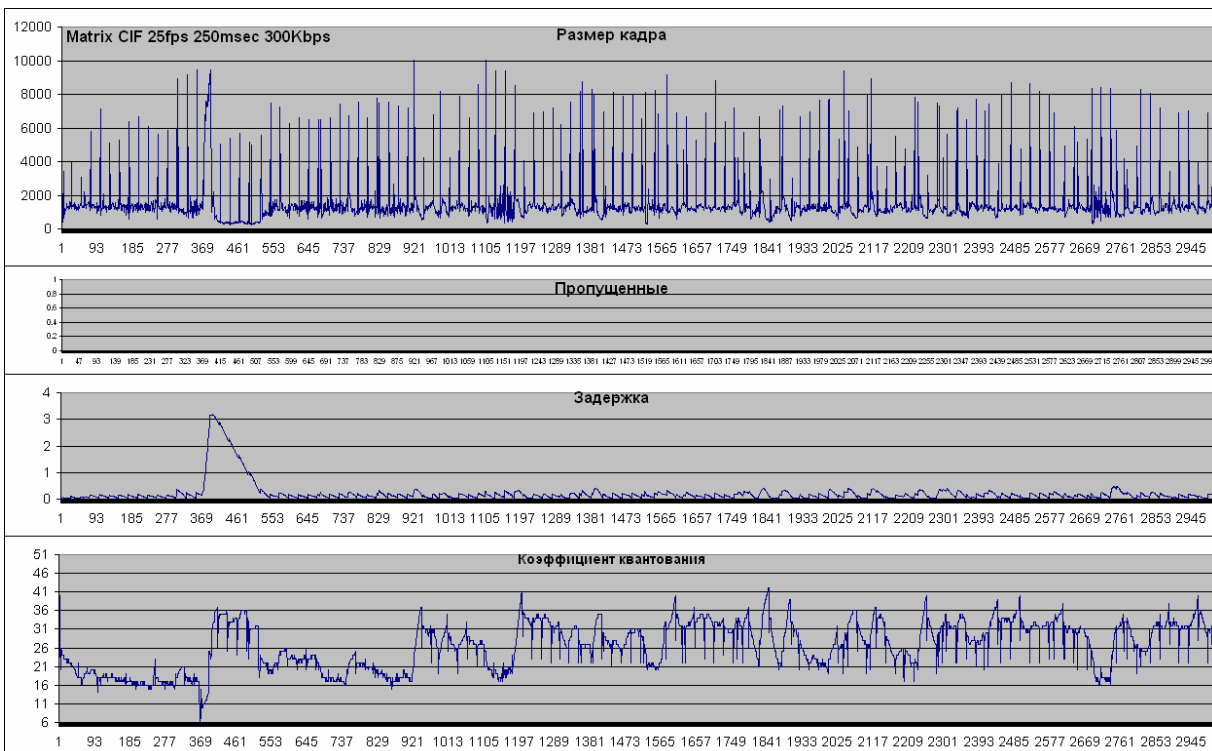


Рис. 1. Кодирование видеопоследовательности без пропуска кадров.

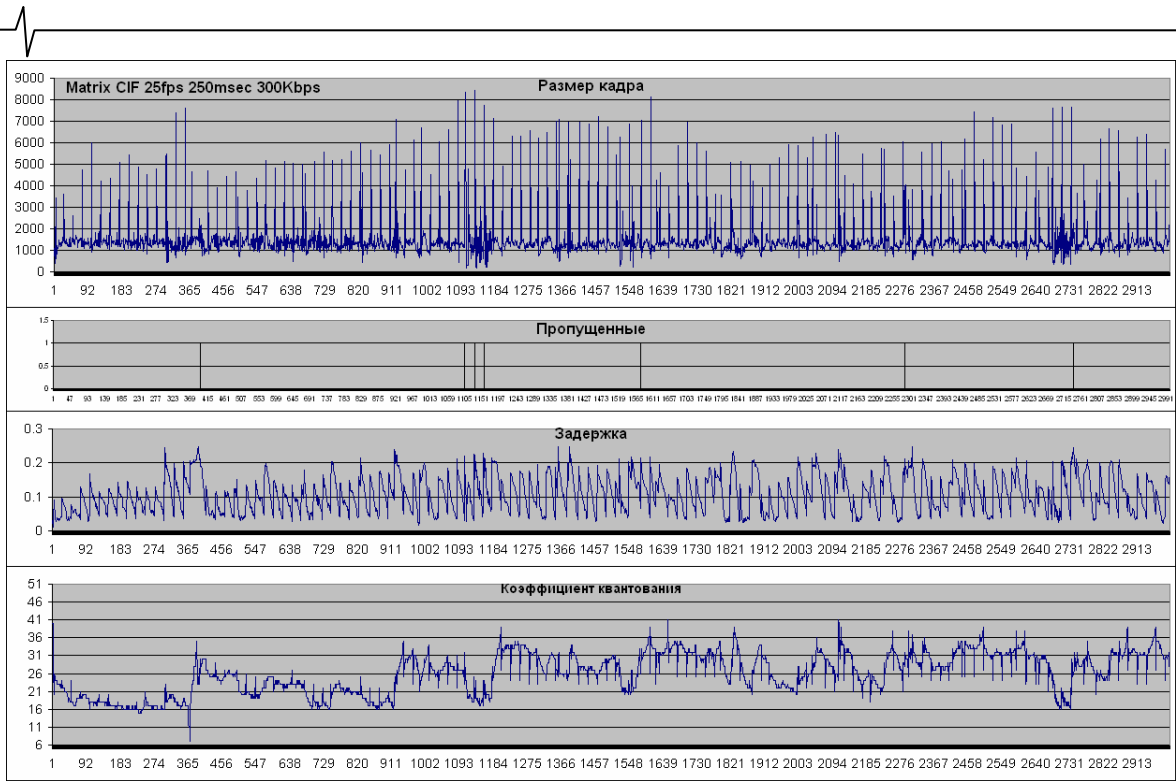


Рис. 2. Кодирование видеопоследовательности с пропуском кадров.



Рис. 3. Ухудшение качества I-кадра при фиксированном соотношении размеров I- и P-кадров (второй кадр – типа I).

Еще одной проблемой является неравномерность качества I- и P-кадров (то есть кадров, закодированных без предсказания движения и с предсказанием движения). Изначально в кодере был реализован вариант, когда в параметрах кодера задается соотношение размеров I- и P-кадров. Проблемой данного подхода является то, что низкое отношение вызывает визуально заметное ухудшение качества I-кадров (см. рис. 3), а высокое – увеличивает вероятность переполнения буфера. Кроме того, для разных пространственных разрешений одно отношение размеров I- и P-кадров может создавать раз-

ные соотношения качества I- и P-кадров.

Поэтому был реализован следующий подход: для I-кадров выбор шаг квантования вычисляется, как коэффициент квантования предыдущего P-кадра, деленный на константу  $r = 1.4$  (пункты 1-2 пересматриваются). Данная константа выбрана эмпирически и позволяет получать I- и P-кадры примерно одинакового качества. При этом учитывается требование размера буфера: из уравнения (1) и условия (3) можно получить минимальный допустимый шаг квантования. Этот подход проиллюстрирован на рис. 4.



Рис. 4. Отсутствие ухудшения качества I-кадра при фиксированном соотношении коэффициентов квантования I- и P-кадров (второй кадр – типа I).

Дальнейшим улучшением алгоритма является перекодирование кадра, если реальный размер закодированного кадра сильно отличается от целевого. Это поможет улучшить качество кодирования и снизить количество пропускаемых кадров.

Данная схема контроля потока и пропуска кадров применима для различных битовых скоростей и временных/пространственных разрешений. Описанный алгоритм контроля потока используется в видеоинформационной системе VPhone. Система обеспечивает высокое качество видеоизображения и малые задержки при интерактивном общении. Достоинства системы особенно проявляются на низкоскоростных каналах связи и каналах связи с жестким ограничением скорости передачи (например, спутниковый канал).

#### Литература

1. ITU-T Recommendation H.264. Advanced Video Coding for Generic Audiovisual Services (ISO/IEC 14496-10. Information Technology – Coding of Audio-Visual Objects – Part 10: Advanced Video Coding). - ITU-T, 2009.
2. Дворкович А.В. Эффективное кодирование видеоинформации в новом стандарте H.264/AVC // Труды НИИП. - М., 2005. - С. 56-69.
3. ITU-T Recommendation H.262, ISO/IEC 13818-2. Information Technology – Generic Coding of Moving Pictures and Associated Audio Information: Video (MPEG-2 Video). - ITU-T, 2000.
4. ISO/IEC 14496-2 – Information Technology – Coding of Audio-Visual Objects – Part 2: Visual (MPEG-4 Video). - ISO, 2004.
5. ITU-T Recommendation H.263. Video Coding for Low Bit Rate Communication. - ITU-T, 2005.
6. H.-J. Lee, T. Chiang, Y.-Q. Zhang. Scalable rate control for MPEG-4 video. // IEEE Trans. on Circuits and Systems for Video Technol. – Sep. 2000, v. 10, № 6, pp. 878-894.
7. T. Chiang, Y.-Q. Zhang. A new rate control scheme using a new rate-distortion model. // IEEE Trans. on Circuits and Systems for Video Technol. – Feb. 1997, v. 7, № 1, pp. 246–250.
8. Дворкович А.В., Быстров С.В. Методика регулировки потока при сжатии динамических видеоизображений // Цифровая обработка сигналов. - 2003. - № 2. - С. 12-16.