

Проблемы и перспективы развития систем кодирования динамических изображений

Продолжение. Начало в №№2,3, 4/2011 *Виктор Дворкович, Михаил Чобану*

В первых трех частях статьи рассматривались основные тенденции развития систем сжатия видеoinформации, основные этапы стандартизации видеопереработки и большинство стандартов компрессии. В данной части речь идет о стандарте H.264/AVC.

H.264/AVC является новейшим международным стандартом видеокодирования, разработанным совместно VCEG и MPEG. В отличие от MPEG-4 Visual, имеющего много вариантов функциональности, H.264/AVC возвращается к традиционной цели – эффективному сжатию и устойчивости к сетевым средам для универсального видеоконтента и для кадра прямоугольной формы. Полное описание применения H.264/AVC приведено в [35]. H.264 позволил добиться существенного улучшения сжатия и качества, которое некоторые тесты оценивают как в 4 раза лучшее объективное качество, чем у MPEG-4 Part 2. Это позволяет получить видео стандартной четкости вещательного качества при потоке в 1,5 Мбит/с, что примерно соответствует 12 сжатым телевизионным программам, передаваемым в полосе частот, прежде занятой одним аналоговым ТВ-каналом. H.264 можно использовать для сжатия фильма высокой четкости и размещения его на простом DVD, или осуществить потоковую передачу видео качества VHS через низкоскоростные каналы ADSL со скоростью 600 Кбит/с.

Проект H.264/AVC был запущен VCEG в начале 1998 года и первоначально обозначался H.26L. Предусматривалось удвоение эффективности кодирования по сравнению с любыми другими существующими стандартами и его использование для разнообразных приложений. В 2001 году был принят начальный вариант разработки H.26L. В 2003 году была завершена первая фаза разработки стандарта, направленная на формальное одобрение, принятое позже ITU-T под именем H.264 и ISO/IEC как MPEG-4, Part 10 (AVC – улучшенное видеокодирование) в комплекте стандартов MPEG-4. В первой версии стандарта использовался только цветовой формат 4:2:0 и разрешение в 8 бит/отсчет. При этом

не поддерживалось высокое разрешение видеосигналов, используемое в наиболее развитых профессиональных средах. Чтобы соответствовать потребностям профессиональных приложений, было запущено продолжение совместного проекта с целью добавления новых возможностей, названных расширениями диапазона точности (fidelity range extensions – FRExt) к исходному стандарту. Проект FRExt, заверченный в 2005 году, содержал комплект из четырех новых профилей, названных «высокими профилями». В 2007 году набор из высоких профилей был увеличен до восьми. Общий обзор H.264/AVC FRExt приведен в [36]. В последние годы были разработаны две поправки к H.264/AVC – масштабируемое видеокодирование (Scalable video coding – SVC) и многокадрное видеокодирование (multiview video coding – MVC). Их рождение обусловлено увеличением потребностей новых приложений, а также ростом сетевых возможностей.

Проект **Video Coding Layer (VCL) H.264/AVC** по существу является схемой гибридного кодирования, но значительно превосходит по быстродействию все предыдущие стандарты. Выигрыш получен за счет более точного предсказания и более эффективного энтропийного кодирования, а также множества новых элементов кодирования. H.264/AVC поддерживает гибкую структуру слоев (slice). Это означает, что размеры и формы слоев не обязательно кратны числу строк макроблоков, как это было в предыдущих стандартах. Кодированное изображение может состоять из различных типов слоев. H.264/AVC поддерживает ссылки на закодированные ранее мультиизображения. Управление изображением обеспечивает высокую степень гибкости в выборе ссылочного изображения, которое ограничивается только максимальной емкостью общей памяти.

Стандарт видеокодирования H.264/AVC предназначен для технических решений, включающих, по крайней мере, следующие области применения:

◆ кабельное, спутниковое, наземное вещание, вещание с помощью кабельных и DSL-модемов;

- ◆ хранение на оптических и магнитных носителях, DVD и т.д.;
- ◆ диалоговые службы, работающие в различных сетях (например, ISDN, Ethernet, LAN, DSL, беспроводные и мобильные сети, а также любые сочетания этих сетей);
- ◆ «видео по запросу» или службы потокового мультимедиа в различных сетях;
- ◆ MMS (службы мультимедийных сообщений) в различных сетях и т.д.

Более того, новые приложения могут развертываться в существующих и перспективных сетях. Стандарт включает 16 профилей (рис. 4) и 16 уровней, ориентированных на конкретные классы приложений и определенные варианты используемых форматов динамических изображений.

Функциональные возможности H.264/AVC позволяют поднять эффективность компрессии по отношению к предшествующим стандартам кодирования видео. Следует подчеркнуть следующие усовершенствования и новые возможности предсказания содержания кодируемого изображения:

- ◆ *компенсация движения с использованием переменных размеров блока, включая малые размеры блока.* Стандарт обладает достаточно большой гибкостью в выборе размеров и форм блоков при компенсации движения по сравнению со всеми предшествующими стандартами, причем минимальный размер блока составляет 4×4 пиксела и позволяет точнее выделять области движения;
- ◆ *компенсация движения с точностью до четверти пиксела.* Стандарт повышает точность компенсации движения до четверти пиксела (что особенно актуально для медленного движения). Цветность, как правило, имеет разрешение, уменьшенное вдвое по вертикали и горизонтали, и компенсация движения для компонент цветности имеет точность в одну восьмую пиксела цветности;
- ◆ *векторы движения, выводящие за границы изображения.* Методика экстраполяции за границы изображения, появившаяся как опция в H.263, включена также в стандарт H.264/AVC;

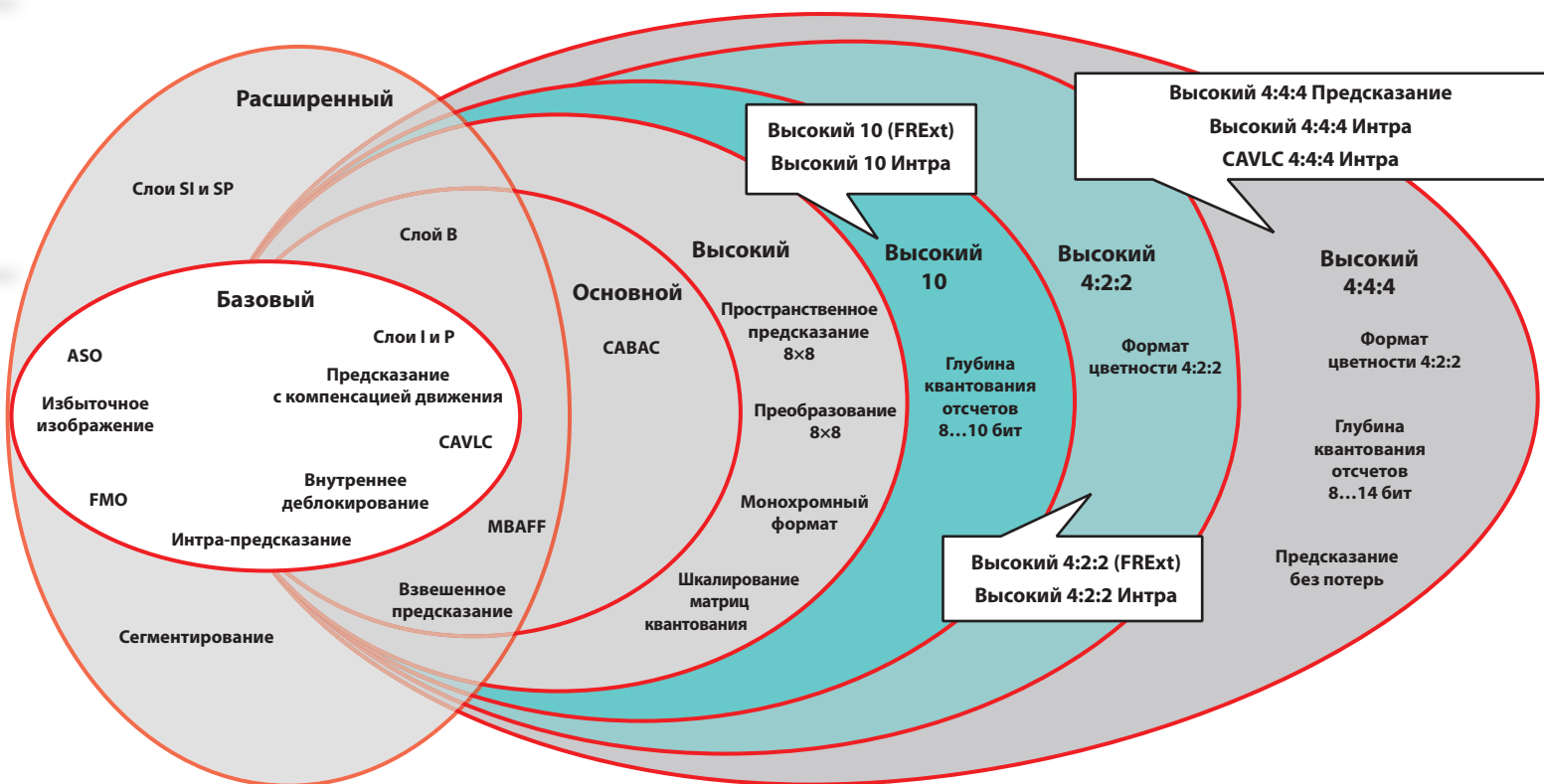


Рис. 4. Профили H.264/AVC [37]

- ◆ **компенсация движения с несколькими опорными изображениями.** В стандарте расширяется методика увеличенного выбора опорного изображения, что поднимает эффективность кодирования, так как позволяет кодеру для компенсации движения делать выбор между большим количеством изображений, декодированных и сохраненных в декодере. Такое расширение возможностей выбора опорного изображения применяется как для P-, так и для B-кадров. Предусмотрена возможность использования до 32 ссылок на другие кадры. В большинстве сцен данная функция обеспечивает не очень значительное повышение качества и не дает заметного понижения цифрового потока. Однако для некоторых сцен, например, с часто повторяющимися участками, возвратно-поступательным движением и т.п., подобный подход при сохранении качества позволяет очень сильно снизить затраты цифрового потока;
- ◆ **независимость порядка воспроизведения изображений и порядка опорных изображений.** В стандарте жесткая зависимость между порядком следования изображений в значительной мере устранена, что позволяет кодеру выбирать порядок изображений для компенсации движения и для воспроизведения с высокой степенью гибкости, ограниченной только объемом памяти, который гарантирует возможность декодирования;
- ◆ **независимость методов обработки изображений и возможности их использования для предсказания движения.** Стандарт обеспечивает кодеру во многих случаях возможность использовать для предсказания движения изображение, близкое по содержанию к кодируемому;
- ◆ **взвешенное предсказание.** Эта новая возможность стандарта позволяет взвешивать и сдвигать сигнал после компенсации движения на величины, указанные кодером. Такая методика может чрезвычайно сильно поднять эффективность кодирования для сцен с изменением освещенности, а также гибко использоваться для других целей;
- ◆ **улучшенная обработка пропущенных (skipped) блоков и блоков с прямым (direct) предсказанием движения.** Стандарт предполагает возможное наличие движения в пропущенных областях. Для областей с двунаправленным предсказанием предусматривается расширенный метод компенсации движения, известный как прямая компенсация движения, которая позволяет улучшить прямое предсказание;
- ◆ **направленное пространственное предсказание для внутрикадрового кодирования.** Новая методика экстраполяции краев ранее декодированных частей текущего изображения применяется в областях изображений, кодируемых с помощью внутрикадровой методики (то есть без ссылок на содержание других изображений). Эта методика повышает качество сигнала, используемого для предсказания, а также позволяет применять для предсказания соседние области, которые были закодированы не с помощью внутрикадровой методики;
- ◆ **гибкие функции чересстрочного сжатия (поддерживается не во всех профилях).** Адаптивное к изображению кодирование полей, позволяющее кодировать каждый кадр как кадр или как пару полей (полукадров) – в зависимости от отсутствия/наличия движения. Адаптивное к макроблокам кодирование полей позволяет независимо кодировать каждую вертикальную пару макроблоков (блок 16×32) как прогрессивные или чересстрочные;
- ◆ **квантование.** Логарифмическое управление длиной шага для упрощения распределения цифрового потока кодером и упрощенного вычисления обратной длины квантования. Частотно-оптимизированные матрицы масштабирования квантования, выбираемые кодером для оптимизации квантования на основе человеческих особенностей восприятия (поддерживается не во всех профилях);
- ◆ **деблокировочная фильтрация в цикле кодирования.** Применение адаптивного деблокировочного фильтра – хорошо известный метод повышения качества восстановленного видео, и при правильном проектировании может повысить как субъективное, так и объективное

качество. Деблокинг-фильтр в стандарте H.264/AVC внесен в петлю обратной связи предсказания и компенсации движения.

Стандарт **H.264/AVC** предусматривает ряд дополнительных функций обработки динамических изображений:

- ◆ **преобразование блоков небольшого размера.** Стандарт основан главным образом на преобразовании блоков 4×4. Это позволяет кодеру представлять сигнал более локально-адаптивным образом, что уменьшает искажения, известные под названием «окантовки» (ringing);
- ◆ **иерархическое преобразование блоков.** Хотя в большинстве случаев использование малых размеров блоков преобразования (4×4) визуально предпочтительнее, есть некоторые сигналы, содержащие существенную корреляцию, для которых выгоднее использовать более длительные базисные функции. Стандарт позволяет делать это двумя способами: с помощью иерархического увеличения эффективного размера блоков цветности до 8×8 пикселей и блоков яркости до 16×16 пикселей;
- ◆ **преобразование с использованием 16-битной арифметики.** Стандарт использует только 16-битную точность вычислений;
- ◆ **точное обратное преобразование.** Стандарт, следуя пути, намеченному в одном из приложений к H.263, первым достиг точного совпадения декодированного видео на всех декодерах;
- ◆ **арифметическое энтропийное кодирование.** В стандарт включен передовой метод энтропийного кодирования, известный как арифметическое кодирование. Он использует эффективную методику – контекстно-адаптивное двоичное арифметическое кодирование CABAC (Context-adaptive binary arithmetic coding). Поддерживается только в Главном профиле и выше. Предусмотрено также использование контекстно-адаптивного кодирования кодами переменной длины – CAVLC (Context-adaptive variable-length coding). CAVLC имеет меньшую сложность, чем CABAC, но, тем не менее, сложнее и эффективнее, чем алгоритмы, применяемые для тех же целей в более ранних технологиях сжатия видео (как правило, это был алгоритм Хаффмана).

Устойчивость к ошибкам и потерям данных и гибкость работы на множестве сетевых структур в стандарте определяется целым рядом новых методик, в числе которых:

- ◆ **структура набора параметров.** Набор параметров обеспечивает устойчивую и эффективную передачу заголовочной информации. Так как потеря нескольких ключевых бит информации может оказать сильное отрицательное воздействие на процесс декодирования, эта ключевая информация отделена и обрабатывается более гибким специализированным способом;
- ◆ **синтаксическая структура блока Уровня сетевой абстракции (NAL unit).** Каждая синтаксическая структура в стандарте помещается в логический пакет данных, называемый блоком NAL. Синтаксическая структура блоков NAL предусматривает достаточно большую гибкость для передачи видеоконтента способом, наиболее подходящим для каждого специфического вида сетей;
- ◆ **гибкий размер слоя.** Слоевая структура задается очень гибко (как это было ранее в стандарте MPEG-1), уменьшая при этом объем заголовочной информации;
- ◆ **гибкое упорядочивание макроблоков.** Используется новая возможность деления изображения на области, называемые «группами слоев», причем каждый слой становится независимо декодируемым подмножеством группы слоев. Гибкое упорядочивание макроблоков может существенно повысить устойчивость к потере данных путем обработки пространственных зависимостей между областями, которые кодируются в каждом слое;
- ◆ **произвольное упорядочивание слоев.** Так как каждый слой кодированного изображения может быть декодирован независимо от других слоев изображения, стандарт позволяет посылать и получать слои в произвольном порядке друг относительно друга. Эта возможность позволяет снизить задержку в приложениях реального времени, особенно при использовании в сетях, имеющих режим работы «доставка вне очереди» (например, в IP-сетях);
- ◆ **избыточные изображения.** Для увеличения устойчивости к потере данных стандарт содержит возможность послышки кодером избыточного представления областей изображений, позволяя воспроизвести области изображений, для представления которых данные были потеряны в процессе передачи;
- ◆ **разбиение данных.** Так как некоторая кодированная информация для представления каждой области (например, вектора движения и другая информация предсказания) имеет большую значимость для представления видеоконтента, чем остальная информация, новый

стандарт позволяет разделить синтаксис каждого слоя на части (до трех частей) для передачи в зависимости от категории синтаксических элементов;

- ◆ **SP/SI-синхронизация/переключение изображений.** Стандарт вводит новые типы изображений, которые дают возможность точной синхронизации процесса декодирования некоторых декодеров с непрерывным видеопотоком. Это свойство может позволить переключение работы декодера между представлениями видеоконтента с разными скоростями потока, восстановление после потерь или ошибок данных, а также использование специальных режимов, таких как быстрая перемотка вперед или назад и др.

Уровень сетевой абстракции (NAL) создан для обеспечения «дружественности к сетям», позволяя простую и эффективную настройку к использованию уровня видеокодирования (VCL) в широком спектре систем. NAL облегчает возможность переносить данные уровня видеокодирования H.264/AVC на транспортном уровне, таком, как:

- ◆ RTP/IP для любых типов проводных и беспроводных интернет-сервисов в реальном масштабе времени (диалоговых или потоковых);
- ◆ файловые форматы, например, ISO, MP4 для хранения и MMS;
- ◆ H.32X для проводных и беспроводных диалоговых сервисов;
- ◆ систем MPEG-2 для вещательных услуг и т.д.

Полная настройка видеоконтента для нужд каждого конкретного приложения находится вне области стандартизации H.264/AVC, но дизайн уровня сетевой абстракции превосходит множество таких настроек.

Сравнение основных применяемых стандартов кодирования приведено в [38].

Литература

35. Я. Ричардсон. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения. М.: Техносфера, 2005, 369 с.
36. Sullivan J., Topiwala P., Luthra A. The H.264/AVC advanced video coding standard: overview and introduction to the fidelity range extensions. In: SPIE Conference on Applications of Digital Image Processing XXVII (August 2004)
37. Pereira F. Recent Achievements on Video Coding. VISNET Summer School, KOC University, Istanbul, Turkey, 2009.
38. <http://www-ee.uta.edu/Dip/Courses/EE5359/standardsreview>

Окончание следует