

## НОВЫЕ ТЕХНОЛОГИИ ВИДЕОКОДИРОВАНИЯ В СТАНДАРТЕ H.264

Дворкович А.В.

### НИИР

Стандарт видеокодирования MPEG-2 (известный также как Рекомендация МСЭ-Т H.262), который был разработан около 10 лет назад и расширил возможностей обработки видео предшествовавшего ему стандарта MPEG-1 (например, поддержка чересстрочного кодирования), явился технологией, позволившей развивать системы цифрового телевидения в общемировом масштабе. Он широко используется для передачи телевизионных сигналов стандартного разрешения и телевидения высокой четкости по спутниковым, кабельным и наземным каналам связи, а так же для хранения видео на DVD.

Однако возрастающие потребности в услугах по доставке видео по каналам связи с различной пропускной способностью создают огромную потребность в повышении эффективности видеокодирования. Рост эффективности кодирования может позволить передавать больше видеоканалов или повысить качество воспроизводимого видеоизображения в рамках существующих пропускных способностей цифровых каналов связи.

Развитие методики кодирования видео для телекоммуникационных приложений прошло через разработку стандартов и рекомендаций МСЭ-Т H.261, H.262 (MPEG-2), H.263, компрессированное видео распространялось по каналам ISDN, T1/E1, телефонной сети общего пользования, мобильным беспроводным сетям и сетям Интранет/Интернет. На протяжении всего этого развития не прекращались попытки увеличить эффективность кодирования при передаче данных по разнообразным типам сетей, при различных требованиях к уровню ошибок и потере данных.

Не так давно стандарт кодирования видео MPEG-4 стал достаточно интенсивно использоваться в области приложений, где применялись предшествующие стандарты. Он обеспечил возможность кодирования формы видеообъектов, он также способствовал расширению области применения цифрового видео.

В начале 1998 года VCEG (Группа Экспертов по Видеокодированию, МСЭ-Т SG16 Q.6) объявила о приеме предложений по проекту H.26L, целью которого было поставлено удвоение эффективности кодирования, то есть уменьшение в два раза скорости выходного потока кодера при заданном качестве воспроизводимого изображения, по сравнению со всеми другими уже существующими стандартами видеокодирования для широкого диапазона приложений. Первый проект нового стандарта был принят в октябре 1999 г. В декабре 2001 г. VCEG и MPEG (Экспертная Группа по Движущимся Изображениям, ISO/IEC JTC 1/SC 29/WG 11) сформировали Совместную Видеогруппу (JVT) с полномочиями завершения проекта нового стандарта видеокодирования для формального утверждения как H.264/AVC в марте 2003 г.

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

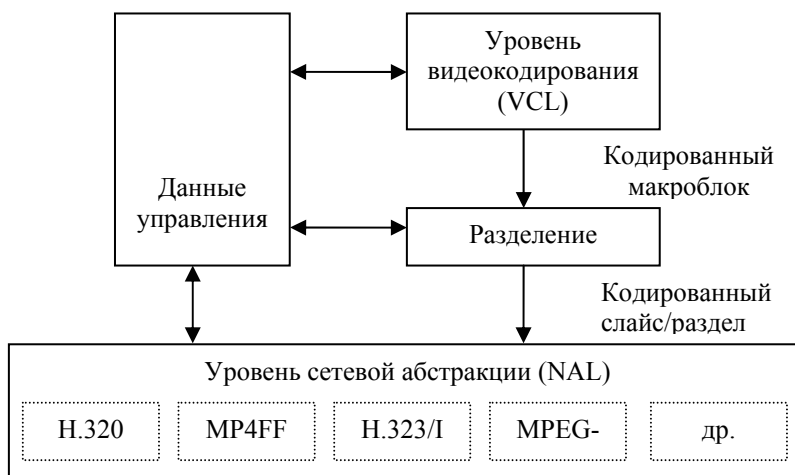
- кабельное, спутниковое, наземное телевизионное вещание;
- хранение на оптических и магнитных носителях, DVD и т.д.;
- интерактивные службы, работающие с видеоинформацией в сетях ISDN, Ethernet, DSL, беспроводных и мобильных сетях и т.д., а также при любых сочетаниях этих сетей;
- видео по запросу и службы потокового мультимедиа в различных сетях;
- службы мультимедийных сообщений в различных сетях.

Новые приложения могут развертываться в существующих и вновь создаваемых сетях.

Для того чтобы удовлетворить потребности в гибкости и настраиваемости, H.264/AVC включает два уровня:

- Уровень видеокодирования (Video Coding Layer, VCL), предназначенный для эффективного представления видеоконтента, и
- Уровень сетевой абстракции (Network Abstraction Layer, NAL), который форматирует представление видео на уровне видеокодирования и обеспечивает заголовочную информацию в виде, подходящем для передачи и хранения.

Ниже приведена структура видеокодера H.264/AVC.



Среди отличительных черт, которые позволяют поднять эффективность кодирования по отношению к предшествующим стандартам кодирования видео, следует выделить следующие улучшения возможностей предсказания содержания кодируемого изображения:

- *Компенсация движения с переменным размером блока, включая малые размеры блока.* Стандарт поддерживает большую гибкость в выборе размеров блоков при компенсации движения, причем минимальный размер блока составляет 4x4 пиксела.

- *Компенсация движения с точностью до четверти пиксела.* Большинство предшествовавших стандартов поддерживали компенсацию движения с точностью не более полупиксела. Новый стандарт повышает точность компенсации движения до четверти пиксела, как это было сделано в расширенном профиле MPEG-4 (в новом стандарте уменьшена сложность интерполяции).

- *Вектора движения, выводящие блоки за границы изображения.* В MPEG-2 и предшествовавших ему стандартах вектора движения могли указывать только на пиксела, находящиеся в границах декодированного опорного изображения. Методика экстраполяции за границы изображения, появившаяся как опция в H.263, включена и в новый стандарт.

- *Компенсация движения с несколькими опорными изображениями.* Изображения, закодированные с предсказанием (P), в MPEG-2 и предшествовавших ему стандартах использовали только одно предыдущее изображение для предсказания значений в новом изображении. В новом стандарте расширяются возможности выбора опорного изображения (такая возможность впервые появилась в приложении к стандарту H.263); это поднимает эффективность кодирования, так как позволяет кодеру выбирать для компенсации движения блоки из большего количества изображений, декодированных и сохраненных на декодере. Такое же расширение возможностей выбора опорного изображения применяется так же и для двунаправленного предсказания. В MPEG-2 для двунаправленного предсказания должны использоваться строго определенные изображения – предшествующий и последующий в порядке воспроизведения внутрикодированные (I) или закодированные с предсказанием (P).

- *Независимость порядка воспроизведения изображений и порядка опорных изображений.* В предшествовавших стандартах устанавливалась жесткая зависимость между порядком следования изображений для использования при компенсации движения и порядком следования изображений на воспроизведении. В новом стандарте эти ограничения в значительной мере устранены, что позволяет при кодировании выбирать порядок изображений для компенсации движения и для воспроизведения с высокой степенью гибкости, которая ограничена только объемом памяти, гарантирующим возможность декодирования. Устранение ограничения также позволяет устранить дополнительную задержку, ранее связанную с двунаправленным предсказанием.

- *Независимость методов обработки изображений и возможности использования их для предсказания движения.* В предшествовавших стандартах изображения, закодированные с использованием некоторых методов (например, двунаправленного предсказания), не могли служить в качестве опорных для предсказания движения других изображений видеопоследовательности. Устраняя это ограничение, новый стандарт обеспечивает кодеру большую гибкость и, во многих случаях, возможность использовать для предсказания движения более близкое к кодируемому (по содержанию) изображение.

- *Взвешенное предсказание.* Эта новая возможность стандарта позволяет взвешивать и сдвигать сигнал после компенсации движения на величины, указанные кодером, что может чрезвычайно сильно поднять эффективность кодирования для сцен с изменением освещенности, а также гибко использоваться для других целей.

- Улучшенная обработка «пропущенных» (*skipped*) блоков и «прямого» (*direct*) предсказания движения. В предшествовавших стандартах «пропущенные» области изображения, кодируемого с предсказанием, не могли двигаться. Это могло приводить к искажениям, когда кодируемое видео содержало движение сцены в целом. Новый стандарт предполагает движение в «пропущенных» областях. Для областей с двунаправленным предсказанием (так называемых В-слайсов) новый стандарт также включает расширенный метод компенсации движения, известный под названием «прямая» компенсация движения, который позволяет улучшить «прямое» предсказание, введенное в приложении к H.263 и MPEG-4.

- Направленное пространственное предсказание при внутрикадровом кодировании. Новая методика экстраполяции границ 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 использовать для предсказания соседние области, которые были закодированы не с помощью внутрикадровой методики (это не допускалось в стандартах H.263 и MPEG-4).

- Деблокинговая фильтрация в цикле кодирования. Кодирование видео, основанное на обработке блоков, приводит к искажениям, называемым «блокингом». Причиной может служить как предсказание, так и кодирование остатков предсказания. Применение адаптивного деблокингового фильтра – хорошо известный метод повышения качества восстановленного видео, и при правильном проектировании он может повысить как субъективное, так и объективное качество. Построенный на основе концепции, заложенной в приложении к стандарту H.263, деблокинговый фильтр в новом стандарте внесен в цикл обратной связи предсказания и компенсации движения. Таким образом, повышение качества изображения может быть использовано в межкадровом предсказании, что, в свою очередь, улучшает возможность предсказания движения других изображений.

Помимо улучшения методов предсказания, другие части стандарта были также расширены с целью повышения эффективности кодирования, а именно:

- Преобразование блоков небольшого размера. Все основные предшествовавшие стандарты кодирования видео использовали косинусное преобразование блоков размером 8x8, в то время как новый стандарт основан главным образом на преобразовании блоков 4x4. Это позволяет кодеру представлять сигнал более локально-адаптивным образом, что уменьшает искажения, известные под названием «окантовки» (*ringing*). Меньший размер блоков также частично обусловлен тем, что использование описанных выше методов позволяет лучше предсказывать содержание блоков, но при этом требуется обеспечить области преобразования с границами, соответствующими меньшим блокам предсказания.

- Иерархическое преобразование блоков. Хотя в большинстве случаев использование малых размеров блоков преобразования (4x4) визуально предпочтительнее, есть некоторые сигналы, содержащие существенную корреляцию, для которых выгоднее использовать более длительные базисные функции. Новый стандарт позволяет делать это двумя способами: (1) с помощью иерархического преобразования (для увеличения эффективного размера блоков) для низкочастотной цветностной информации (8x8); и (2) путем предоставления кодеру возможности выбрать специальный тип внутрикадрового кодирования, позволяющий расширить размер преобразования яркости для низкочастотной информации до размера блока 16x16 способом, очень похожим на тот, что применяется для цветности.

- Преобразование с использованием 16-битной арифметики. Все предшествовавшие стандарты требовали достаточно сложных вычислений при преобразовании как на кодере, так и на декодере. В то время как ранее требовались, как правило, вычисления с использованием 32-битной арифметики, новый стандарт требует только 16-битной точности вычислений.

- Точное обратное преобразование. В предшествовавших стандартах видеокодирования преобразование, используемое при кодировании видео, оговаривалось с учетом допустимых ошибок из-за практической нереализуемости точного обратного преобразования. В итоге каждая реализация декодера могла давать слегка различное декодированное видео. Новый стандарт, следуя пути, намеченному в одном из приложений к H.263, реализует точное совпадение декодированного видео на всех декодерах.

- Арифметическое энтропийное кодирование. В стандарт включен передовой метод энтропийного кодирования, известный как арифметическое кодирование. Хотя арифметическое кодирование включалось как необязательная возможность в стандарте H.263, новый стандарт использует более эффективную методику, называемую САВАС (контекстно-адаптивное двоичное арифметическое кодирование).

- Контекстно-адаптивное энтропийное кодирование. В новом стандарте применяются два метода энтропийного кодирования – САVLC (контекстно-адаптивное кодирование кодами переменной длины) и САВАС (контекстно-адаптивное двоичное арифметическое кодирование). Адаптивность, основанная на контексте (уже закодированных данных), повышает производительность кодирования по сравнению с методиками в предыдущих стандартах.

Устойчивость к ошибкам и потерям данных и гибкость работы на множестве типов сетей в новом стандарте определяется внедрением ряда новых методик, в том числе следующих:

- Структура набора параметров. Набор параметров обеспечивает устойчивую и эффективную передачу заголовочной информации. Так как потеря нескольких ключевых бит информации (таких как заголовок последовательности или заголовок изображения) может оказать сильное отрицательное

воздействие на процесс декодирования в предыдущих стандартах, эта ключевая информация была отделена, и обрабатывается более гибко в новом стандарте.

- *Синтаксическая структура блока Уровня сетевой абстракции (NAL unit)*. Каждая синтаксическая структура в новом стандарте помещается в логический пакет данных, называемый блоком NAL. Вместо создания специфического потокового интерфейса системы (как в предшествовавших стандартах видеокодирования) синтаксическая структура блоков NAL дает большую гибкость передачи видеоконтента способом, наиболее подходящим для каждого специфического вида сетей.

- *Гибкий размер слайса*. В отличие от жесткой слайсовой структуры MPEG-2 (которая уменьшает эффективность кодирования, увеличивая объем заголовочной информации и уменьшая эффективность предсказания), размеры слайсов в новом стандарте задаются очень гибко (как это было ранее в стандарте MPEG-1).

- *Гибкое упорядочивание макроблоков*. Была разработана новая возможность деления изображения на области, называемые группами слайсов, причем каждый слайс становится независимо декодируемым подмножеством группы слайсов. При эффективном использовании гибкое упорядочивание макроблоков может существенно повысить устойчивость к потере данных путем обработки пространственных зависимостей между областями, которые кодируются в каждом слайсе. Гибкое упорядочивание макроблоков может также использоваться для множества других целей.

- *Произвольное упорядочивание слайсов*. Так как каждый слайс закодированного изображения может быть декодирован независимо от других слайсов изображения, новый стандарт позволяет посылать и получать слайсы в произвольном порядке друг относительно друга. Эта возможность, впервые появившаяся в необязательном приложении к стандарту H.263, может снизить задержку в приложениях реального времени, особенно при использовании на сетях, имеющих режим работы «доставка вне очереди» (например, IP-сетях).

- *Избыточные изображения*. Для увеличения устойчивости к потере данных новый стандарт содержит новое свойство, позволяющее кодеру посылать избыточные представления областей изображений, позволяя воспроизвести области изображений (обычно с некоторой потерей качества), для представления которых данные были потеряны в процессе передачи.

- *Разбиение данных*. Так как некоторая закодированная информация для представления каждого региона (например, вектора движения и другая информация предсказания) имеет большую значимость для представления видеоконтента, чем остальная информация, новый стандарт позволяет разделить синтаксис каждого слайса на части (до трех частей) для передачи в зависимости от категории синтаксических элементов. Эта часть стандарта построена на основе, заложенной в MPEG-4 и необязательном приложении к H.263. В новом стандарте методика упрощена благодаря использованию единого синтаксиса, который делится на соответствующие разделы, а управление синтаксисом производится с помощью введения определенной категоризации его элементов.

- *SP/SI синхронизация/переключение изображений*. Новый стандарт включает новое свойство, заключающееся во введении типов изображений, которые делают возможной точную синхронизацию процесса декодирования некоторых декодеров с непрерывным видеопотоком, производимым другими декодерами, не приводя к потере эффективности декодеров, обусловленной посылкой внутрикодированного (I) изображения. Это свойство может обеспечить возможность переключения режимов 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.

#### Литература

1. Joint Video Team of ITU-T and ISO/IEC JTC 1, “Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC)”, Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVT-G050, March 2003.
2. ITU-T and ISO/IEC JTC 1, “Generic coding of moving pictures and associated audio information – Part 2: Video”, ITU-T Recommendation H.262 – ISO/IEC 13818-2 (MPEG-2), November 1994.
3. ITU-T, “Video Codec for Audiovisual Services at px64 kbit/s”, ITU-T Recommendation H.261, Version 1: November 1990; Version 2: Mar. 1993.
4. ITU-T, “Video coding for low bit rate communication”, ITUT Recommendation H.263; version 1, November 1995; version 2, Jan. 1998; version 3, Nov. 2000.
5. ISO/IEC JTC1, “Coding of audio-visual objects – Part 2: Visual”, ISO/IEC 14496-2 (MPEG-4 visual version 1), April 1999; Amendment 1 (version 2), February, 2000; Amendment 4 (streaming profile), January, 2001.
6. T.Wiegang, G.J.Sullivan, G.Bjontegaard, and A.Luthra, “Overview of the H.264/AVC Video Coding Standard”, IEEE Transactions on Circuits and Systems for Video Technology, July 2003

