

**«Современные методы и стандарты
экономного кодирования
видеоинформации»**

**Семенюк В. В.
(E-mail: ecoding@mail.ru)**

**Санкт-Петербург
2002**

Введение

Не так давно разработанный стандарт MPEG4 уже успел получить достаточно широкое распространение. На сегодняшний день имеется множество примеров его практической реализации, причем как в программных приложениях, так и на аппаратном уровне. Можно констатировать, таким образом, что новый стандарт оказался достойной заменой более раннему стандарту MPEG2. Тем не менее прогресс не стоит на месте, и уже сегодня стоит обратить самое пристальное внимание на новые перспективные технологии, которые рано или поздно могут прийти на смену существующему стандарту. Данная статья содержит обзор существующих и новых, еще не получивших распространение методов экономного кодирования видеoinформации. Автор сознательно ограничивается рассмотрением способов кодирования видеосигнала, оставляя за рамками обзора методики обработки аудиoinформации.

Первая часть обзора содержит краткое описание методов кодирования информации, используемых в существующем стандарте MPEG4. Дальнейший обзор посвящен двум новым технологиям, скрывающимся под аббревиатурами H.264 (более раннее название – H.26L; стандарт, пришедший на замену стандарту H.263, а также являющийся официальным расширением стандарта MPEG4) и MVC (технология кодирования, предлагаемая компанией Nokia; описывается версия 2.0).

Стандарт MPEG4

Стандарт MPEG4 базируется на более раннем стандарте MPEG2, который в свою очередь является модификацией стандарта MPEG1. Как следствие, многие методы, используемые в стандарте MPEG4, уже были реализованы в более ранних стандартах.

Предполагается, что исходная видеопоследовательность представлена в формате YUV 4:2:0 и имеет горизонтальные и вертикальные размеры изображения, кратные 16 (наиболее часто используемые размеры – 352x288 (так называемый *формат CIF*)). При таком представлении изображение хранится в виде трех цветовых компонентов: одного компонента яркости (Y) и двух компонентов цветовой насыщенности (U, V). Размеры яркостного компонента совпадают с размерами исходного изображения, тогда как размеры компонентов цветовой насыщенности в два раза меньше по каждому из измерений. Данное представление очень удобно, так как позволяет существенно облегчить процесс обработки информации за счет концентрации большей части обрабатываемой информации в одном из компонентов (Y). Возможность такой концентрации обуславливается особенностями зрительного восприятия человека – человеческий глаз более чувствителен к изменению яркости и менее чувствителен к изменению цвета.

При обработке видеопоследовательности каждое изображение (фрейм) может быть представлено в виде суперпозиции *объектных видеоплоскостей* (*video object plane - VOP*). Максимальное количество таких плоскостей – 256. Каждая плоскость также представляет собой изображение в формате YUV 4:2:0. Она содержит определенную часть исходного (полного) изображения и, как правило, формально связана с какой-то визуально различимой его частью. Плоскости кодируются независимо. При декодировании изображение «собирается» из отдельных декодированных плоскостей с использованием хранимой отдельно информации о форме (*shape*) объектов, содержащихся в той или иной плоскости (информация о форме образует так называемые *альфа-плоскости* (*alpha planes*)). В частном случае мы имеем всего одну плоскость, которая содержит все исходное изображение. Вследствие отсутствия эффективных методов разделения изображения на отдельные объекты, именно этот случай является наиболее часто встречающимся.

При обработке видеоплоскости подразделяются на четыре типа: I-плоскость, P-плоскость, B-плоскость и S-плоскость. I-плоскости кодируются без использования какой-либо информации о других плоскостях. I-плоскость можно декодировать, не декодируя никакие другие плоскости, что позволяет обеспечить быстрый доступ к произвольным частям закодированного видеопотока. Каждая I-плоскость разбивается на *макроблоки* размера 16x16 (здесь и далее указывается размер для компонента Y; соответствующий размер для компонентов U и V – 8x8). Такие макроблоки называются *intra-макроблоками*. Intra-макроблок в свою очередь разбивается на 4 intra-блока размера 8x8, каждый из которых участвует в последующем *блочном* кодировании/декодировании (все сказанное относится только к компоненту Y; для компонентов U и V в качестве кодируемого блока выступает блок 8x8, соответствующий макроблоку компонента Y).

P-плоскости кодируются с использованием информации о предшествующих I- и P-плоскостях. P-плоскость также разбивается на макроблоки 16x16 и блоки 8x8, однако, помимо intra-макроблоков и intra-блоков, здесь присутствуют *inter-макроблоки* и *inter-блоки*. Перед кодированием над inter-макроблоками/блоками осуществляется процедура *компенсации движения* (*motion compensation*): в предшествующих плоскостях ищется макроблок/блок (он может находиться на произвольной позиции, не обязательно кратной 16 или 8), максимально соответствующий текущему inter-макроблоку/блоку (т.н. *предсказание* (*prediction*)), после чего макроблоки/блоки поточечно вычитаются друг из друга; получившийся в результате этой операции разностный макроблок/блок (*ошибка предсказания* (*prediction error*)) в дальнейшем используется при кодировании. Важно отметить, что плоскости, на основе которых осуществляется процедура компенсации движения (*опорные плоскости* (*reference planes*)), должны совпадать с плоскостями, восстанавливаемыми в процессе декодирования (при вычислении ошибки предсказания ни в коем случае не должна использоваться информация, доступная только на этапе кодирования). Результатом действия описанной процедуры помимо блоков, содержащих ошибку предсказания, является поле двухкомпонентных *векторов движения* (*motion vectors*).

Поиск схожих блоков производится только для компонента Y, вектора для компонентов U и V получаются путем деления пополам векторов для первого компонента. Согласно стандарту, вектора движения имеют в общем случае нецелую длину, кратную 1/2 или 1/4. Для получения нецелых векторов, опорные плоскости увеличиваются (масштабируются) соответственно в 2 и в 4 раза по каждому из измерений (для увеличения используется билинейная интерполяция). Помимо увеличения, стандарт предписывает применение так называемого *пэддинга* (*padding*). Процедура заключается в расширении плоскости со всех сторон на величину, равную 16 точкам для компонента Y и 8 – для компонентов U и V (и то и другое в целых единицах). Новым точкам присваиваются цвета ближайших к ним граничных точек исходной плоскости. Пэддинг (в данном случае, возможно, имеет смысл использовать термин *дополнение/расширение*) способствует повышению эффективности компенсации движения вблизи границ видеоплоскостей.

Компенсация движения может осуществляться как макроблоками, так и блоками, поэтому для каждого макроблока могут быть найдены либо 1, либо 4 вектора движения. Выбор осуществляется на основе сравнения объемов, которые будет занимать закодированный макроблок в том и в другом случае с учетом вклада кода вектора движения.

Как было отмечено, P-плоскости могут содержать как inter-, так и intra-макроблоки. Благодаря компенсации движения, inter-макроблоки, как правило, кодируются с большей эффективностью по сравнению с intra-макроблоками, однако использование intra-макроблоков не связано с хранением дополнительной информации о векторах движения, что опять же может оказаться более выгодным. Дилемма разрешается так же, как и в предыдущем случае: тип используемого макроблока выбирается исходя из объема его суммарного кода.

Стандарт предусматривает особый тип компенсации – *компенсация с перекрытием (overlapped motion compensation)*. Компенсация с перекрытием используется только для блоков компонента Y. При использовании этого метода разность берется не между текущим блоком и некоторым схожим блоком, принадлежащим предыдущим плоскостям, а между текущим блоком и взвешенной суперпозицией трех схожих блоков. В качестве соответствующих трех векторов движения выбираются следующие вектора: вектор для данного блока и два вектора для блоков, соседних к данному блоку в обрабатываемом макроблоке.

В-плоскость есть расширение понятия P-плоскости. Отличие В-плоскости от P-плоскости заключается в том, что при кодировании В-плоскости для компенсации движения могут использоваться как предыдущие, так и последующие I- и P- плоскости (по понятным причинам, сами В-плоскости для этого использоваться не могут). Соответственно вводятся понятия *прямое* и *обратное предсказание (forward and backward prediction)*. Каждый макроблок в В-плоскости может быть предсказан макроблоком на предыдущей плоскости, макроблоком на последующей плоскости и суперпозицией этих макроблоков. Не вдаваясь в подробности процедуры компенсации движения для В-плоскостей, отметим, что использование *двухнаправленного предсказания (bidirectional prediction)* позволяет заметно повысить эффективность предсказания, а следовательно, и эффективность кодирования.

Последний тип закодированной видеоплоскости – S-плоскость – имеет непосредственное отношение к *спрайтам (sprite)*. Спрайтом в стандарте MPEG4 называется часть изображения, видимая на протяжении определенного интервала в видеопоследовательности. В роли спрайта чаще всего выступает фон (задний план). При декодировании с использованием спрайтов фреймы частично восстанавливаются из отдельных областей спрайта путем отображения этих областей на ту или иную область декодируемого фрейма с использованием перспективного преобразования (помимо отображения, применяются также и другие специальные процедуры). Спрайт обычно хранится и кодируется отдельно, при этом способ кодирования идентичен способу кодирования I-плоскостей.

Частным случаем использования спрайтов являются S(GMC)-плоскости. В роли спрайта здесь выступает одна из предшествующих I- или P-плоскостей. Этот метод почти ничем не отличается от метода обработки P-плоскостей, за тем исключением, что процедура компенсации движения предваряется процедурой глобальной (применяемой ко всей плоскости) компенсации движения с использованием перспективного преобразования. Технология *GMC (global motion compensation)*, чаще всего применяется для нивелирования движения и изменения оптических параметров камеры.

Идущие подряд плоскости различных типов делятся на обособленно кодируемые группы. В начале группы должна находиться I-плоскость, так как только она может быть декодирована первой. Далее расположены плоскости других типов. В-плоскости обычно чередуются с P- или S(GMC)-плоскостями (хотя это и не обязательно; например, нередко одна P- или S(GMC)-плоскость чередуется с двумя и более идущими подряд В-плоскостями). Деление плоскостей на группы облегчает процесс доступа к информации, а также обеспечивает большую помехоустойчивость. Последнее обусловливается тем, что в случае ошибки поврежденным оказывается не весь закодированный поток, а только отдельная его часть.

Рассмотрим вкратце процедуры кодирования intra/inter-блоков и векторов движения. Intra- и inter-блоки кодируются практически одним и тем же методом за небольшими исключениями: первоначально производится блочное *дискретное косинусное преобразование (discrete cosine transform - DCT)*, затем полученные в результате преобразования коэффициенты (*DCT-коэффициенты*) квантуются, после чего квантованные коэффициенты кодируются с использованием заранее фиксированной системы кодов переменной длины. Основное отличие метода кодирования intra-блоков от метода кодирования inter-блоков заключается в том, что в intra-блоках по-особому обрабатывается старший коэффициент преобразования, описывающий средний по блоку уровень сигнала (*DC-коэффициент*): кодируется не сам коэф-

фициент, а разность между ним и DC-коэффициентом некоторого соседнего блока. Схема кодирования всех остальных коэффициентов выглядит следующим образом. Осуществляется обход блока в зигзагообразном порядке – от старшего коэффициента, находящегося в левом верхнем углу блока, к младшему, находящемуся в правом нижнем углу (предусмотрены три варианта зигзагообразного обхода). В процессе обхода последовательно кодируются группы коэффициентов следующего типа: последовательность нулевых коэффициентов и первый следующий за ними ненулевой коэффициент. Группы нумеруются, причем один номер резервируется для специфического случая – ситуации, когда все оставшиеся коэффициенты являются нулевыми. Групповым номерам сопоставляются коды переменной целой длины. Длина выбирается исходя из следующего критерия: номерам групп, встречающимся чаще, ставятся в соответствии более короткие коды. Система кодов является *префиксной*: коды, принадлежащие системе, не являются началом других кодов из этой системы. Префиксность делает возможным однозначное декодирование.

Способ кодирования блоков коэффициентов, полученных в результате применения дискретного косинусного преобразования, обусловлен следующей особенностью: при обработке реальных изображений, как правило, DCT-коэффициенты убывают по абсолютной величине при движении по блоку слева направо и сверху вниз; соответственно, во время зигзагообразного обхода коэффициенты убывают по абсолютной величине, причем, будучи квантованными, очень многие (особенно коэффициенты, оказывающиеся при обходе последними) имеют нулевое значение.

Помимо приведенного способа кодирования, существует несколько модифицированный способ, при котором из коэффициентов верхней строки и левого столбца блока предварительно вычитаются коэффициенты, стоящие на тех же позициях в соседних блоках.

Стандартом предусматривается две схемы квантования коэффициентов, получаемых в результате дискретного косинусного преобразования. Одна из них повторяет схему, используемую в стандарте H.263, и заключается в делении всех коэффициентов, кроме DC-коэффициента в *intra*-блоках, на одно и то же число – удвоенный параметр квантования. При этом в случае *inter*-блоков перед квантованием из абсолютного значения коэффициента вычитается значение, равное половине параметра квантования. DC-коэффициент в *intra*-блоке всегда делится на 8. В альтернативной схеме более тщательно учитываются особенности зрительного восприятия информации. Для DC-коэффициента используется неравномерное квантование, при котором делитель зависит не только от параметра квантования, но и от типа цветового компонента. Все остальные коэффициенты кодируются в два этапа. Сначала каждый из них делится на соответствующее ему число из стандартной для данного типа блока (*intra* или *inter*) матрицы квантования (размер матрицы совпадает с размером блока – 8x8). Матрица устроена таким образом, чтобы при движении слева направо и сверху вниз делители увеличивались. (Дело в том, что зрение сильнее воспринимает изменения, описываемые низкими частотами гармонического спектра, поэтому высокими частотами часто можно практически безболезненно пренебречь.) На втором этапе частично квантованные коэффициенты повторно делятся на число, равное удвоенному параметру квантования.

Вектора движения кодируются по координатно. Как и в случае с DCT-коэффициентами, при генерации кода используются системы префиксных кодов переменной длины. Кодируется не сама координата вектора, а разность между ней и одной из соответствующих координат вектора некоторого соседнего блока (в выборе участвуют координаты векторов трех соседних блоков; выбирается средняя из трех координат). Различным значениям разности координат сопоставляются различные коды переменной длины, при этом маленьким по абсолютному значению разностям соответствуют более короткие коды.

Описанные методы кодирования коэффициентов и векторов движения задействуются не всегда: возникают ситуации, когда данные процедуры вообще не требуются. Одной из таких ситуаций является случай равенства нулю всех DCT-коэффициентов блока. О том, выполня-

ется это условие или нет, сигнализирует специальный флаг. Помимо этого флага, существует также и другой флаг, наличие которого позволяет упростить процедуру кодирования. Он сигнализирует о том, кодируется ли данный *inter*-макроблок вообще. Если флаг имеет единичное значение, никакая дополнительная информация, описывающая коэффициенты и вектора движения не передается: в качестве векторов движения берутся нулевые вектора, также нулевой полагается ошибка предсказания.

В стандарте оговариваются два метода пост-обработки видеопоследовательности: *деблокинг* (*deblocking*) и *дерингинг* (*deringing*). Первый применяется для уменьшения блочного эффекта – особого типа ошибки восстановления закодированного изображения, обусловленного поблочным кодированием с высоким уровнем потерь; второй представляет собой технику адаптивного сглаживания сигнала, позволяющую устранять отдельные высокочастотные искажения.

Среди прочего стандарт MPEG4 описывает формат представления некоторых весьма специфических визуальных объектов: *неподвижное изображение* (*still image*), *лицевая анимация* (*face animation*) и *сетчатый объект* (*mesh object*).

Неподвижное изображение (как правило, обычная картинка) кодируется с применением вейвлет-преобразования. Осуществляется стандартная двухмерная частотная декомпозиция, при этом в качестве базиса может быть использован либо ставший уже классическим в вейвлет-технологиях биортогональный базис (9,3), либо некий специфический базис, параметры которого заранее неизвестны и передаются вместе с кодом. Коэффициенты, получаемые в результате вейвлет-преобразования, квантуются с использованием равномерного скалярного квантователя, а затем кодируются по стандартной схеме, в которой применяются так называемые *нуль-деревья* (*zero-tree*). Код генерируется на основе высокоэффективного *арифметического кодирования*.

В стандарте описывается достаточно оригинальная технология, предназначенная для описания мимики человеческого лица. Лицо восстанавливается на основе его параметрической модели (лицевая анимация). Не вдаваясь в подробности, отметим, что унифицированное описание позволяет представлять лицо в достаточно экономной форме по сравнению с обычными методами кодирования (например, описанными выше). Наиболее вероятная область применения такого решения - кодирование видеоконференций.

Специальное представление объектов на плоскости в виде треугольной сетчатой структуры с наложенной на нее текстурой носит название *mesh-технология*. Описание топологии сетки осуществляется путем кодирования координат вершин треугольников, а также векторов движения, определяющих изменение положения этих вершин во времени. Параллельно с трансформацией топологии сетки производится трансформация накладываемой текстуры. Сетчатое представление объектов оказывается эффективно в тех случаях, когда эти объекты изначально задаются в векторном виде.

Выше кратко изложены методики обработки информации, применяемые в стандарте MPEG4. Перейдем к рассмотрению основных составляющих двух новых технологий кодирования.

Стандарт H.264

При разработке стандарта впервые были объединены усилия специалистов Международного Телекоммуникационного Союза (International Telecommunication Union - ITU) в лице их стандартизационного сектора (ITU-T), Международной Организации по Стандартизации (International Organization for Standardization – ISO) и Международной Комиссии по Электротехнике (International Electrotechnical Commission – IEC). До объединения первая организа-

ция занималась разработкой телекоммуникационных стандартов, тогда как последние две – стандартов общего плана. Ранее ими были предложены стандарты кодирования видеoinформации H.261, H.263 (ITU-T); JPEG, JPEG2000, MPEG (ISO/IEC).

В отличие от стандарта MPEG4, технологии, описываемые в стандарте H.264, предназначены для обработки обычных видеопоследовательностей, предварительно не разделенных на видеоплоскости. Естественно, это не является ограничением, так как все методы кодирования могут быть с тем же успехом применены и к случаю многоплоскостного кодирования.

Описываемые в новом стандарте методы в целом не сильно отличаются от методов, предусмотренных стандартом MPEG4. Тем не менее новый стандарт включает в себя несколько достаточно перспективных решений, заслуживающих самого пристального внимания.

В качестве базовых единиц кодирования в данном случае выступают блоки размера 4x4, 8x8, 8x4 и 4x8. При этом размер макроблока остается тем же – 16x16, что дает возможность осуществлять компенсацию движения с большим количеством различных вариантов разбиения макроблока на блоки. Разбиение осуществляется в два этапа. Сначала макроблок 16x16 разбивается на одну или нескольких прямоугольных областей, каждая из которых имеет свой, отличный от других вектор движения (возможны следующие варианты разбиения: 16x16, 16x8+16x8, 8x16 + 8x16 и 8x8 + 8x8 + 8x8 + 8x8). На втором этапе каждый подблок размера 8x8 в свою очередь разбивается на области одним из нижеперечисленных способов: 8x8, 8x4+8x4, 4x8+4x8 и 4x4 + 4x4 + 4x4 + 4x4.

Для блоков 4x4 применяется весьма специфическое преобразование, отличное от дискретного косинусного преобразования. Согласно утверждениям авторов, данное преобразование обладают схожей с дискретным косинусным преобразованием эффективностью, но меньшей вычислительной сложностью (используются только операции сложения и сдвига). Весьма оригинальной идеей в данном случае является повторное кодирование старших коэффициентов преобразования. Кодирование каждого блока представляет собой повторное применение 24 процедур кодирования блоков (16 для компонента Y и по 4 для компонентов U и V). Полученные 16 старших коэффициентов преобразования для компонента Y и, соответственно, по 4 старших коэффициента для компонентов U и V подвергаются повторному преобразованию. Для блока старших коэффициентов компонента Y используется несколько отличное от предыдущего преобразование 4x4, а для блока старших коэффициентов компонентов U и V – специальное преобразование 2x2.

Для блоков 8x8, 8x4 и 4x8 также предусмотрены свои преобразования, которые, однако, не являются обязательными. Вместо кодирования на основе преобразования может применяться экстраполяция. При использовании такого решения значения яркости/цвета внутри блока получаются путем приближения, в котором учитываются значения яркости/цвета граничных точек соседних блоков.

Как видно, компенсация движения претерпела довольно существенные изменения по сравнению со стандартом MPEG4. Следует выделить также еще две его особенности: большая точность представления векторов движения и другой способ увеличения масштаба опорных фреймов, используемых для получения предсказания. Допускается точность, равная либо 1/4, либо 1/8 (в MPEG4 максимальная точность, как известно, составляет 1/4). Для масштабирования опорных фреймов применяется достаточно сложная многоточечная интерполяция. В совокупности данные решения позволяют получить довольно весомый прирост в эффективности.

Большой интерес вызывает метод кодирования квантованных коэффициентов преобразования. Для intra-макроблоков/блоков применяется так называемое *intra-предсказание*. Из значений коэффициентов intra-макроблоков/блоков вычитаются числа, получаемые из совокупностей значений некоторых коэффициентов соседних макроблоков/блоков с использованием одной из нескольких схем предсказания (для макроблоков предусмотрено 4 схемы, а для блоков – 9 схем). При кодировании полученных разностей для случая intra-блоков и оши-

бок предсказания для случая *inter*-блоков задействуется метод, практически идентичный методу, описываемому в стандарте MPEG4.

Действенным приемом увеличения эффективности представления видеoinформации оказывается использование для генерации кода арифметического кодирования взамен префиксного. Особенностью предлагаемой реализации арифметического кодирования является то, что кодирование принимает на вход только бинарные символы. С одной стороны, это несколько затрудняет применение данного способа генерации кода (приходится использовать бинарную декомпозицию для многосимвольных алфавитов), с другой стороны, подобный подход делает возможными достаточно быстрые реализации и несколько облегчает процесс построения вероятностных моделей. В стандарте описывается обладающая высокой производительностью табличная реализация метода.

Стандарт H.264 изначально ориентирован на применение, прежде всего, в коммуникационных системах, поэтому особое внимание здесь уделено повышению помехоустойчивости и обеспечению высокой степени гибкости информационного представления, необходимой для обеспечения удобства передачи закодированной видеoinформации по сетям разной природы. Предусматриваются особые алгоритмы разделения закодированного потока на блоки, обеспечивающие требуемые характеристики информационного представления для каждого конкретного случая.

В заключение рассмотрения стандарта H.264 приведем его краткую характеристику. Стандарт действительно является заметным шагом вперед по сравнению с текущей версией стандарта MPEG4 (как было указано, рассмотренный стандарт должен стать расширением данного стандарта). Формат кодирования позволяет добиваться эффективности представления информации, превосходящей эффективность, допустимую в рамках стандарта MPEG4, более чем на 30%. Естественно, вычислительная сложность предлагаемых методов несколько выше вычислительной сложности стандартных на сегодняшний день решений, однако, по мнению автора данного обзора, это обстоятельство не должно стать препятствием для внедрения нового стандарта.

Технология MVC

Рассматриваемая в данном разделе технология кодирования вряд ли имеет перспективы стать международным стандартом, однако некоторые весьма оригинальные идеи, лежащие в ее основе, без сомнения, заслуживают упоминания.

Подобно стандарту H.264, технология MVC ориентирована на применение к последовательностям фреймов, а не видеоплоскостей, что, как уже было указано, ни в коей мере не ограничивает ее общности.

Общая схема алгоритма обработки информации, так же как и в случае со стандартом H.264, практически идентична схеме, описанной в стандарте MPEG4. Отличия, опять же, заключаются в деталях реализации.

Разбиение макроблока на блоки очень напоминает метод, используемый в MPEG4. Основное отличие заключается в том, что, если макроблок в компоненте *Y* разбивается на 4 блока 8x8, соответствующий ему макроблок 8x8 в компонентах *U* и *V* разбивается на 4 блока 4x4.

Для получения базисного разложения задействуется несколько пространственно-частотных преобразований. Помимо стандартного дискретного косинусного преобразования 8x8, предусматривается дискретное косинусное преобразование для блоков 8x4, 4x8 и 4x4, а также альтернативное преобразование для блоков 8x8 и 4x4, некорректно именуемое *преобразованием Кархунена-Лоева (Karhunen-Loeve transform)*. Альтернативное преобразование

позволяет хорошо описывать цветовые возмущения, распространяющиеся фронтально от некоторой наклонной линии, проходящей через центр блока. Взамен базисного разложения может применяться экстраполяция аналогично тому, как это делается в стандарте H.264.

Так же как и в стандарте H.264, для intra-блоков используется intra-предсказание. Предусмотрено в общей сложности 13 методов предсказания, выбор которых осуществляется по достаточно сложной схеме.

Основной интерес в технологии MVC привлекает методика компенсации движения. Для получения предсказания блока используется 6-параметрическое аффинное преобразование. При вычислении ошибки из блока вычитается прямоугольная область, аффинным преобразованием переводимая в некоторую область фрейма, используемого для предсказания. Для вычисления точных параметров преобразования задействуются бикубические сплайны для компонента Y и билинейная интерполяция для компонентов U и V. При кодировании параметры преобразования квантуются.

В отличие от стандарта H.264, в рассматриваемой технологии для генерации кода не используется арифметическое кодирование, что следует рассматривать как явный недостаток. Однако подбор системы префиксных кодов выглядит вполне оправданным, так что не исключено, что выигрыш от применения арифметического кодирования является все же не столь существенным.

При рассмотрении технологии MVC следует обратить особое внимание на алгоритм деблокинга. Здесь используется достаточно сложная схема нелинейной фильтрации, обеспечивающая очень высокие результаты как по визуальному качеству, так и по количественному отношению сигнал/шум. Тем не менее на данный момент нет смысла подробно останавливаться на рассмотрении данного алгоритма, так как, судя по всему, в дальнейшем он может претерпеть ряд существенных изменений. Следует отметить только, что заметным недостатком предлагаемого решения является его высокая вычислительная сложность.

Как видно, предлагаемая компанией Nokia технология, помимо большого количества стандартных, содержит и ряд новых, достаточно смелых решений. При этом следует отметить, что данная технология является далеко не единственной оригинальной технологией кодирования. На сегодняшний день существует множество аналогичных предложений, и хочется надеяться, что некоторые из них послужат фундаментом для нового поколения технологий и, как следствие, нового поколения стандартов, значительно более эффективных по сравнению с существующими.

Литература

[1] ISO/IEC JTC1/SC29/WG11, *Coding of Moving Pictures and Audio: MPEG-4 Video Verification Model version 18.0*, JTC1/SC29/WG11 N3908, Pisa, January 2001.

[2] Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, *Joint Final Committee Draft (JFCD) of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC)*, JVT-D157, Australia, July 2002.

[3] Nokia Inc., *MVC Decoder Description*, Geneva, February 2000.