

«Обзор стандарта JPEG2000»

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

**Санкт-Петербург
2002**

Введение

В последнее время информационные технологии приобрели особый статус. Если еще несколько десятилетий назад мало кто имел хоть какое-то представление об этой новой и достаточно прогрессивной области знаний, сегодня нелегко найти человека, так или иначе не сталкивавшегося в жизни с ее практическими приложениями. Повсеместная компьютеризация и информатизация, цифровой звук, цифровое видео, цифровая телефония – это лишь отдельные примеры внедрения информационных технологий в быту. По прогнозам специалистов, в ближайшие годы доля подобных приложений будет только увеличиваться. В связи с этим, дальнейшее изучение прикладных аспектов теории информации приобретает особое значение. Некоторые из разрабатываемых технологий уже сегодня выглядят многообещающе. Вероятно, одним из самых ярких примеров этому могут служить методы цифровой обработки видеозображений.

Проблема обработки видеосигнала отнюдь не нова. Уже начиная с середины двадцатого века, ученые занимаются активными исследованиями в данной области. Следствием их усилий на первоначальном этапе стало появление аналогового телевидения. Формирование, передача и отображение движущейся картинки оказываются возможными благодаря применению несложных аналоговых элементов. Практический эффект от подобной новинки сразу привлекает внимание специалистов, что приводит к настоящему телевизионному и киноиндустриальному буму.

Ближе к концу двадцатого века исследователи все чаще направляют свои усилия на разработку цифровых технологий передачи и обработки сигнала. Преимущество цифрового подхода над аналоговым очевидно: цифровое представление более удобно в использовании, оно позволяет добиться большего качества и, подчас, большей вместимости. В конце восьмидесятых перспективность использования цифрового представления уже не вызывает сомнений, а дополнительный толчок к развитию цифровых методов дает появление персональных компьютеров.

Первые персональные вычислительные машины были не достаточно производительны для выполнения сложных задач связанных с обработкой потокового видео. Тем не менее очень часто производительности вполне хватало для осуществления манипуляций с отдельными изображениями. В связи с этим произошло выделение двух составляющих проблемы обработки видео. Проблема обработки потока изображений и проблема обработки одиночных изображений теперь представляли собой две отдельные задачи, причем, имея в своем распоряжении не достаточно быстрые, но, что немаловажно, недорогие персональные вычислительные машины, исследователи имели значительно больше возможностей для решения именно последней задачи. Как следствие, именно обработка неподвижных изображений стала одним из основных приложений персональных компьютеров.

Под обработкой неподвижных изображений обычно подразумеваются не только манипуляции, приводящие к их изменению, но также и операции позволяющие изменять физическое представление изображений без визуально ощутимых последствий. Последнее принято называть *сжатием изображений*.

Процесс сжатия приводит к уменьшению объема представления информации на информационном носителе. Существует два типа сжатия: сжатие, сопряженное с искажениями пред-

ставляемой информации, и сжатие, не допускающее таких искажений. Первое носит название *сжатие с потерями*, а второе – *сжатие без потерь*. По понятным причинам сжатие с потерями применимо далеко не ко всем типам информации, однако там, где оно все таки может быть использовано, оно, как правило, оказывается значительно более эффективным по сравнению со сжатием без потерь. В случае с неподвижными изображениями сжатие с потерями оказалось не только применимым, но и более чем оправданным. Как оказалось, человеческое восприятие не улавливает некоторые визуальные детали; потеря или искажение этих деталей соответственно является совершенно безболезненной для человека.

Исследователями было разработано огромное множество методов сжатия неподвижных изображений. Эти методы обладают различной эффективностью и различной вычислительной сложностью, что позволяет находить решение для очень широкого спектра задач. Однако они не совместимы между собой. Можно воспользоваться любым методом для получения представления изображения, но это не гарантирует возможность корректной интерпретации представления лицом, не вовлеченным в процесс кодирования. Здесь естественным образом встает проблема унификации методов, то есть проблема стандартизации.

Координацией усилий по стандартизации методов сжатия видеоизображений традиционно занимается международная организация по стандартизации (ISO) в лице специальной группы экспертов по фотографическим изображениям (JPEG), а также отделение стандартизации международного телекоммуникационного союза (ITU-T). Разработка первого полноценного стандарта, начатая в 1986 году, увенчалась появлением в 1992 году международного стандарта JPEG. Формирование стандарта проводилось на конкурсной основе. Среди множества предложений было выбрано одно, которое было взято за основу стандарта; в дальнейшем предложение было усовершенствовано усилиями различных специалистов.

JPEG в целом оказался достаточно удачным стандартом. Он обеспечивает высокую эффективность сжатия при приемлемом уровне потерь и допускает эффективные практические реализации. К недостаткам стандарта традиционно относят функциональную ограниченность формата данных – форма представления информации препятствует ее использованию в приложениях, связанных с видеоредактированием.

Стандарт JPEG широко используется на практике с самого появления. Он стал одним из наиболее популярных стандартов хранения информации. На сегодняшний день стандарт поддерживает практически каждое приложение, ориентированное на обработку видеоизображений.

Несмотря на огромную популярность стандарта JPEG, по прошествии нескольких лет естественным образом возникла необходимость в появлении нового стандарта. К этому времени были разработаны более эффективные методы сжатия, а приложения требовали от форматов хранения видеоданных все больше и больше функциональности. Стандарт JPEG часто не удовлетворял новым требованиям и должен был быть заменен.

Название нового стандарта – JPEG2000 – изначально указывало на год его ожидаемого выхода. К сожалению, на сегодняшний день данный стандарт так и окончательно не увидел свет, хотя основные этапы его утверждения все же были успешно пройдены (в частности, JPEG2000 приобрел официальный статус международного стандарта). Мы с нетерпением ожидаем скорого появления полноценной версии стандарта и предлагаем его краткое описа-

ние. Материалом для данной статьи стали официальные предварительные версии нового стандарта, а также следующие англоязычные обзоры:

[1] Adams M. D., *The JPEG-2000 Still Image Compression Standard*, ISO/IEC JTC 1/SC 29/WG 1 N 2412, Sept. 2001.

[2] Marcellin M. W., Gormish M. J., Bilgin A. and Boliek M. P., *An Overview of JPEG-2000* (invited paper), in Proc. of 2000 Data Compression Conference, pp. 523-541, Snowbird, Utah, March 2000.

[3] Marcellin M. W., Bilgin A., *JPEG2000: Highly Scalable Image Compression* (invited paper), in Proc. of 2001 International Conference on Information Technology: Coding and Computing (ITCC2001), pp. 268-272, Las Vegas, Nevada, April 2001.

Стандарт JPEG2000

Отправной точкой для стандарта JPEG2000 стало предложение М. Болиека 1996 года. Разработанный Болиеким алгоритм должен был стать основой нового стандарта сжатия изображений без потерь JPEG-LS, но был отвергнут в пользу более перспективного алгоритма LOCO-I. Алгоритм Болиека, тем не менее, обладал рядом очень привлекательных возможностей, что послужило причиной создания нового стандарта JPEG2000.

Объявление о начале разработки нового стандарта датируется мартом 1997 года. Традиционно был устроен конкурс алгоритмов сжатия, на котором проводилось численное и визуальное сравнение результатов работы различных программ. Программа-победитель (ею стала разработка арizonского университета, алгоритм WTCQ) была выбрана за основу первой версии стандарта. В ноябре 1998 года с подачи Д. Таубмана в стандарт было внесено существенное изменение. Таубман предложил решение, позволившее сделать стандарт существенно более гибким и менее требовательным к ресурсам вычислительной системы. Алгоритм Таубмана (алгоритм EBCOT) в результате составил основу финальной версии стандарта.

В процессе стандартизации было учтено большое количество различных предложений. Так как все они не могли составить новый стандарт, было принято решение часть из них внести в его базовый вариант, а часть рассматривать как дополнение. На данный момент документация стандарта представлена двумя частями (в будущем планируются дополнительные разделы). Первая описывает основные моменты, которые должны быть в обязательном порядке соблюдены в любой реализации стандарта. Вторая содержит расширения основной части стандарта, которые не являются обязательными. Данный подход выгоден тем, что, во-первых, учитывает большое количество различных предложений и обеспечивает гибкость, а во-вторых, позволяет получать достаточно непритязательные в плане вычислительных ресурсов реализации, совместимые со стандартом.

Разделение стандарта на основную и дополнительную часть лучше всего иллюстрируется на примере алгоритмов квантования. Предложение арizonского университета подразумевало использование сложного алгоритма квантования, получившего название *квантование с сетчатой геометрией*. Фактически была предложена быстрая реализация векторного квантования. Как известно, векторное квантование обладает большей эффективностью по сравнению

со скалярным квантованием, однако, в то же время, оно является существенно менее производительным. В предложенном алгоритме WTCQ векторное квантование реализовано на базе конечного набора скалярных квантователей, выбор которых (для осуществления квантования) производится в соответствии с возможными направлениями обхода заданного графа (сети). Направление обхода одновременно соответствует младшим разрядам квантованных значений, полученных с использованием предыдущих квантователей (выбранных на предыдущем этапе). Путем перебора возможных путей внутри графа можно найти более или менее эффективные способы квантования произвольной по длине последовательности величин. Несмотря на то, что подобный алгоритм весьма прост, его сложность, все же, не идет ни в какое сравнение со сложностью алгоритма обычного равномерного квантования. Как следствие, последний (используется равномерный квантователь с мертвой зоной вблизи нуля) является обязательной частью стандарта, а первый (квантование с сетчатой геометрией) представляет собой расширение стандарта и является лишь его опцией.

Сжатие по стандарту JPEG2000 основано на ставшем уже классическим алгоритме пирамидального вейвлет-преобразования. Обработка вейвлет-коэффициентов осуществляется методом контекстно-зависимого бит-ориентированного арифметического кодирования.

Первоначально изображение подвергается чередующимся последовательностям вертикальных и горизонтальных одномерных вейвлет преобразований. Сначала преобразуются все строки, а затем все столбцы. На следующем этапе левая верхняя четверть матрицы получившейся в результате предыдущего преобразования опять преобразуется (сначала все строки, затем все столбцы). И так далее. Количество этапов соответствует количеству уровней вейвлет-декомпозиции. В результате преобразования мы получаем множество прямоугольных диапазонов вейвлет-коэффициентов, которые принято называть *частотными диапазонами*, так как они содержат информацию о том, как ведет себя исходный двухмерный сигнал (изображение) при разном разрешении (то есть набор коэффициентов при разной частоте).

Для преобразования могут использоваться различные вейвлет-фильтры. Обязательная часть стандарта предписывает использование только двух фильтров: обратимый “5/3” – для сжатия без потерь и необратимый “9/3” – для сжатия с потерями (оба фильтра являются классическими вейвлет-фильтрами). Расширение допускает любые другие фильтры. Подразумевается, что для реализации преобразования используется удобная с практической точки зрения лифтинг-схема.

После преобразования осуществляется квантование коэффициентов. Именно на этапе квантования возникают основные информационные потери, и именно за счет квантования возможно существенное уменьшение объема представления изображения. (Естественно, в квантовании нет необходимости, если производится сжатие без потерь.) Как уже было сказано, квантование может быть либо равномерным скалярным, либо каким-либо другим (например, векторным). В случае использования равномерного скалярного квантования квант-параметр может меняться в зависимости от квантуемого диапазона.

Этап арифметического кодирования является завершающим этапом кодирования. Диапазоны коэффициентов разбиваются на прямоугольные кодовые блоки (как правило, 32x32 или 64x64). Каждый из блоков кодируется независимо. Это означает, что состояние арифметического кодера сбрасывается перед кодированием очередного кодового блока. В процессе кодирования коэффициенты в блоке виртуально представляются в виде битовых плоскостей.

Одну из таких плоскостей составляют знаки коэффициентов; остальные плоскости соответствуют различным разрядам величин коэффициентов (положение бита в плоскости соответствует положению коэффициента в блоке). Кодирование коэффициентов сводится к кодированию битов, составляющих эти коэффициенты. Таким образом, арифметическое кодирование является бит-ориентированным.

Арифметическое кодирование основано на контекстно-зависимой модели. Контекст формируется как функция от значений битов, окружающих кодируемый бит. Кодирование осуществляется по плоскостям: сначала кодируется плоскость, соответствующая старшему разряду коэффициентов, затем следующая по убыванию, и т.д. Во время кодирования каждому коэффициенту в кодируемом блоке ставится в соответствие параметр *значимость*. Коэффициент называется значимым, если в уже закодированных на данный момент битовых плоскостях, присутствует хотя бы один ненулевой разряд данного коэффициента. Каждая битовая плоскость кодируется в три прохода. Во время первого кодового прохода осуществляется распространение информации о значимости коэффициентов. Для каждого бита плоскости, если соответствующий коэффициент еще не является значимым, и если хоть один соседний коэффициент уже является значимым, осуществляется кодирование факта значимости для текущего коэффициента, то есть фактически осуществляется кодирование значения данного бита текущей кодируемой плоскости. Если кодируемый бит оказался ненулевым, сразу после его обработки кодируется соответствующий бит битовой плоскости знаков коэффициентов (кодирование знака). Во время второго кодового прохода кодируются все биты, соответствующие значимым на данный момент коэффициентам и не обработанным во время первого прохода. В отличие от предыдущего кодового прохода, когда решение о кодировании принималось на основе информации о значимости соседних коэффициентов, во время данного прохода биты кодируются в обязательном порядке. Цель третьего кодового прохода – обработать те биты, которые не были обработаны во время первого и второго проходов. Во время третьего прохода арифметическое кодирование применяется совместно с групповым кодированием. Существенной деталью, предусмотренной стандартом, является возможность пропуска кодовых проходов, что является еще одним источником повышения эффективности за счет информационных потерь (первым, наиболее явным источником является квантование). Данная возможность активно используется для осуществления контроля над скоростью генерации кода.

Представление информации, полученное в результате вейвлет-преобразования, очень удобно тем, что оно обеспечивает возможность получения приблизительных копий изображения без осуществления полного обратного преобразования. Обратное преобразование осуществляется в порядке, обратном порядку прямого преобразования. Производя ограниченное число обратных декомпозиций (объединение частотных диапазонов), полагая, что все не вовлеченные в преобразование частотные диапазоны содержат исключительно нулевые элементы, мы легко можем получить либо копию изображения в уменьшенном масштабе, либо исходное изображение, но в более низком качестве по сравнению с изображением, полученным в результате полного обратного преобразования. Учитывая тот факт, что блоки вейвлет-коэффициентов кодируются независимо друг от друга, мы получаем возможность частичного декодирования не только на уровне преобразования, но и на уровне интерпретации кода. Для получения приблизительной копии изображения достаточно декодировать все-

го лишь часть информации, а затем произвести частичное обратное преобразование. Таким образом, формат хранения изображения обеспечивает масштабируемость как по разрешению, так и по качеству.

Другим важным преимуществом нового стандарта является возможность доступа к отдельным элементам изображения без полного декодирования его представления. Обеспечивается такая возможность, во-первых, разбиением исходного изображения на непересекающиеся области (*тайлы*), которые кодируются как отдельные изображения, а во-вторых, представлением кода отдельного тайла в виде частей (*слоев*), каждая из которых является суммарным кодом коэффициентов, соответствующих некоторой его (тайла) области (отметим, что слои в свою очередь делятся на так называемые *пакеты*, содержащие код блоков коэффициентов на разных уровнях декомпозиции). Для того, чтобы декодировать какую-либо область изображения достаточно определить, каким тайлам она принадлежит и какие слои, относящиеся к этим тайлам содержат код блоков коэффициентов, необходимых для восстановления требуемой области.

Безусловно, «удобное» представление изображения не может быть выгодным с точки зрения эффективности сжатия. Действительно, с уменьшением размера структурных элементов (тайлов, областей тайлов, образующих слои и др.) эффективность сжатия несколько снижается. Стандарт в данном случае оставляет нам выбор: с одной стороны, мы имеем возможность получать информационные представления, позволяющие достаточно быстро извлекать и редактировать части изображения, с другой стороны, стандарт не препятствует созданию информационных представлений, эффективных по объему.

Как нетрудно заметить, все, что было сказано выше, в действительности относится не совсем к сжатию изображений. Речь шла всего лишь о сжатии матриц. Реальные изображения подчас являются более сложными объектами. Как правило, изображение включает в себя сразу несколько компонентов. Чаще всего, оно состоит из трех цветовых компонентов: красного, зеленого и синего. Так как каждый компонент в отдельности представляет собой матрицу, для того, чтобы закодировать изображение целиком, нам необходимо закодировать не одну, а три матрицы. Такой подход, как показывает практика, приемлем, но не является самым удачным. Большей эффективности сжатия можно добиться в случае, когда кодируемые компоненты представлены в яркостно-цветовой форме.

Для преобразования изображения из стандартного цветового представления RGB в яркостно-цветовое представление YCrCb стандартом предусмотрены две процедуры: обратимая и необратимая. Необратимая процедура в точности повторяет классическое преобразование RGB-> YCrCb, которое использовалось, например, в старом стандарте JPEG. Обратимая процедура представляет собой достаточно грубое приближение к классической необратимой процедуре. Как следует из названия, данное преобразование не ведет к потере цветовой информации, и может применяться в тех случаях, когда задействуется режим сжатия без потерь.

Для обеспечения помехоустойчивости информационного представления и удобства доступа к информации в стандарте JPEG2000 предусмотрена система маркеров и маркерных сегментов. Маркеры играют роль разграничителей внутри информационного потока; маркерные сегменты содержат в себе параметры фрагментов информации ограниченных маркерами. Данные, начинающиеся с маркера, как правило, могут быть корректно проинтерпрети-

рованы без какой-либо дополнительной информации (это, естественно, не означает возможность восстановления целого по фрагментам), что обеспечивает возможность частичного восстановления изображения, представление которого было повреждено. Введение элементов помехоустойчивости дает зеленый свет использованию стандарта во всевозможных телекоммуникационных приложениях.

В заключение обзора стандарта следует сказать несколько слов относительно его реальной эффективности. Достижение высокого качества сжатия, безусловно, было одной из главных задач при его создании, и здесь разработчики добились явного прогресса. Стандарт JPEG2000 превосходит по эффективности стандарт JPEG примерно в 2 раза при сжатии с потерями и на 5-20% при сжатии без потерь. Конечно, эффективность при сжатии без потерь в данном случае оказывается не такой высокой, как, скажем, у стандарта JPEG-LS, однако она вполне приемлема. Что же касается эффективности сжатия с потерями, здесь стандарт позволяет получать результаты, близкие к наилучшим на сегодняшний день результатам для подобного рода методов.

Судить о скорости работы практических реализаций стандарта пока рано - количество таких реализаций на сегодняшний день очень невелико. Тем не менее можно сказать, что, как и ожидалось, новый стандарт несколько медленнее своего предшественника.

Заключение

JPEG2000, без сомнения, должен рано или поздно стать полноценным стандартом, чему есть все предпосылки. Высокая степень сжатия, возможность быстрого доступа к отдельным частям информации, превосходная структурированность, нетребовательность к ресурсам и много другое – все это будет способствовать повсеместному распространению нового стандарта и, возможно, послужит сильным толчком для дальнейшего развития технологий обработки видеoinформации.

Словарь терминов

arithmetic coding – арифметическое кодирование

bandwidth – полоса пропускания; полоса частот

baseline codec – базовая версия кодека

big endian – порядок следования разрядов от старшего к младшему

binary symbol – двоичный символ

bit-plane coding – кодирование битовых плоскостей

bit-rate (bitrate) – скорость битового потока

bit stream (bitstream) – битовый поток

chrominance component – цветовая компонента

cleanup pass – проход, во время которого осуществляется обработка данных, не обработанных во время предыдущих проходов

code-block – блок коэффициентов, подлежащий кодированию

code-block scan – порядок, в котором кодируются коэффициенты в блоке

coding pass – кодовый проход

compression ratio – степень сжатия

computational complexity – вычислительная сложность

context-dependent coding – контекстно-зависимое кодирование

context label – идентификатор контекста

context model – контекстная модель

DCT (discrete cosine transform) – дискретное косинусное преобразование

DWT (discrete wavelet transform) – дискретное вейвлет-преобразование

dead zone quantization – квантование с мертвой зоной

decomposition level – уровень декомпозиции

delimiting marker – разделяющий маркер

digital imagery – оцифрованные изображения

EBCOT (embedding block coding with optimized truncation) – вложенное блочное кодирование с оптимизированным усечением

EZW (embedded zerotree wavelet) – вейвлет-кодирование на основе вложенных нуль-деревьев

embedding principle – принцип вложенности

entropy coding – кодирование с эффективностью, соответствующей энтропии источника

error resilience – устойчивость к ошибкам

filter bank – совокупность параметров фильтра

functional marker – функциональный маркер

ICT (irreversible component transform) – необратимое преобразование компонентов

ISO (International Organization for Standardization) – международная организация по стандартизации

ITU-T (International Telecommunication Union – Telecommunication Standardization

Sector) – международный телекоммуникационный союз, сектор стандартизации телекоммуникаций

image component – компонента изображения (как правило, цветовая)
image flipping – переверот изображения
image rotation – поворот изображения
irreversible transform – необратимое преобразование
in bit stream marker – маркер, встроенный в битовый поток
inter(-)band dependency – зависимость между диапазонами
inter(-)component transform – межкомпонентное преобразование
interleaving – расслоение; чередование
inverse transformation – обратное преобразование
JPEG (joint photographic experts group) – объединенная группа экспертов по фотоизображениям
L-level dyadic wavelet transform – L-уровневое пирамидальное вейвлет-преобразование
LFS (lowest frequency subband) – диапазон наименьших частот
LSB (least significant bit) – младший разряд
layer – код для области изображения (слой суммарного кода)
lifting-scheme – лифтинг-схема
lossless compression – сжатие без потерь
lossy compression – сжатие с потерями
low/hi bit-rate compression – более/менее высокая степень сжатия
low/hi-pass subband – диапазон нижних/верхних частот
luminance component – яркостная компонента
MSB (most significant bit) – старший разряд
MSE (mean-square error) – среднеквадратическая ошибка
magnitude refinement pass – проход, во время которого осуществляется уточнение разрядов величины
marker segment – маркерный сегмент
near-lossless compression – сжатие с незначительным уровнем потерь
packet – код области на определенном уровне декомпозиции (пакет кода)
packet partition – пакетный сегмент
to pan – панорамировать
precinct – область на определенном уровне декомпозиции
progressive mode – прогрессивный режим
quality scalability – качественная масштабируемость
RCT (reversible component transform) – обратимое преобразование компонентов
ROI (regions of interest) – выделяемые регионы
raster order – растровый (построчный) порядок
rate-distortion – соотношение скорости передачи и уровня искажений
rate control – контроль скорости генерации кода
reconstructed image – реконструированное изображение
reference grid – координатная сетка
resolution level – уровень разрешения
resynchronization marker – маркер повторной синхронизации
run-length encoding – групповое кодирование

SNR (signal-to-noise ratio) – отношение сигнал-шум

SPIHT (set partitioning in hierarchical trees) – кодирование с разделением на множества иерархических деревьев

sample value – значение выборки

scalable compression – сжатие, допускающее масштабируемость

scalar quantization – скалярное квантование

significance propagation pass – проход, во время которого осуществляется обработка информации о значимости разрядов

significance state – значимость

sign-magnitude representation – представление числа в виде <знак, величина>

spatial region – область пространства

spatial scalability – пространственная масштабируемость

source image – исходное изображение

sub(-)band – частотный диапазон

subband decomposition – подразделение на частотные диапазоны

subband recomposition – объединение частотных диапазонов

TCQ (trellis coded quantization) – квантование с сетчатой геометрией

transform coefficient – коэффициент преобразования

tile-component – область пространства внутри изображения

tiling – разбиение пространства на области

VM (verification model) – контрольная модель (для осуществления контроля и сравнения)

variable length codes – коды переменной длины

vector quantization – векторное квантование

visual quality – визуальное качество

WTCQ (wavelet trellis coded quantization) – вейвлет-кодирование с TCQ-квантованием

wavelet filter – вейвлет-фильтр

zero(-)tree – ноль-дерево

to zoom – изменять масштаб