

Сжатие видеоданных

Сжатие видео (англ. **Video compression**) — технология цифровой компрессии телевизионного сигнала, позволяющая сократить количество данных, используемых для представления видеопотока.

Преимущество: Сжатие видео позволяет эффективно уменьшать поток, необходимый для передачи видео по каналам радиовещания, уменьшать пространство, необходимое для хранения данных на носителе.

Недостатки: при использовании сжатия с потерями появляются характерные, иногда отчётливо видимые артефакты — например, блочность (разбиение изображения на блоки 8x8 пикселей), замыливание (потеря мелких деталей изображения) и т. д. Существуют и способы сжатия видео без потерь, но на сегодняшний день они уменьшают данные недостаточно.

Теория

Видео — это по существу трёхмерный массив цветных пикселей. Два измерения означают вертикальное и горизонтальное разрешение кадра, а третье измерение — это время. Кадр — это массив всех пикселей, видимых камерой в данный момент времени, или просто изображение. В видео возможны также так называемые полукадры (см.: чересстрочная развёртка).

Сжатие было бы невозможно, если бы каждый кадр был уникален и расположение пикселей было полностью случайным, но это не так. Поэтому можно сжимать, во-первых, саму картинку — например, фотография голубого неба без солнца фактически сводится к описанию граничных точек и градиента заливки. Во-вторых, можно сжимать похожие соседние кадры. В конечном счёте, алгоритмы сжатия картинок и видео схожи, если рассматривать видео как трёхмерное изображение со временем как третьей координатой.

Сжатие без потерь

Помимо сжатия с потерями видео также можно сжимать и без потерь. Это означает, что при декомпрессии результат будет в точности (бит к биту) соответствовать оригиналу. Однако при сжатии без потерь невозможно достигнуть высоких коэффициентов сжатия на реальном (не искусственном) видео. По этой причине практически всё широко используемое видео является сжатым с потерями (в том числе на потребительских цифровых видеодисках, видеохостингах, в спутниковом вещании). На веб-сайтах для маленьких роликов без звука иногда используются простые форматы GIF и APNG.

Сжатие видео и технология компенсации движения

Одна из наиболее мощных технологий, позволяющая повысить степень сжатия, — это компенсация движения. При любой современной системе сжатия видео последующие кадры в потоке используют похожесть областей в предыдущих кадрах для увеличения степени сжатия. Однако, из-за движения каких-либо объектов в кадре (или самой камеры) использование подобия соседних кадров было неполным. Технология компенсации движения позволяет находить похожие участки, даже если они сдвинуты относительно предыдущего кадра.

Современное состояние дел

На конец 2011 года практически все алгоритмы сжатия видео (например, стандарты, принятые ITU-T или ISO) используют дискретное косинусное преобразование (DCT) или

его модификации для устранения пространственной избыточности. Другие методы, такие как фрактальное сжатие и дискретное вейвлет-преобразование, также были объектами исследований, но сейчас обычно используются только для компрессии неподвижных изображений.

Использование большинства методов сжатия (таких, как дискретное косинусное преобразование и вейвлет-преобразование) влечёт также использование процесса квантования. Квантование может быть как скалярным, так и векторным, тем не менее, большинство схем сжатия на практике используют скалярное квантование вследствие его простоты.

При использовании традиционных методов сохранения информации электронная версия фильма получится слишком большой. Достаточно очевидное усовершенствование состоит в том, чтобы первый кадр запомнить целиком (в литературе его принято называть ключевым), а в следующих сохранять лишь отличия от начального кадра (разностные кадры).

Принцип формирования разностного кадра поясняется рис. 2, где продемонстрировано небольшое горизонтальное смещение прямоугольного объекта. Отчетливо видно, что при этом на всей площади кадра изменились всего 2 небольшие зоны: первая сзади объекта возвратилась к цвету фона, а на второй - перед ним, фон перекрасился в цвет объекта. Для разноцветных предметов произвольной формы эффект сохранится, хотя изобразить его будет заметно труднее.

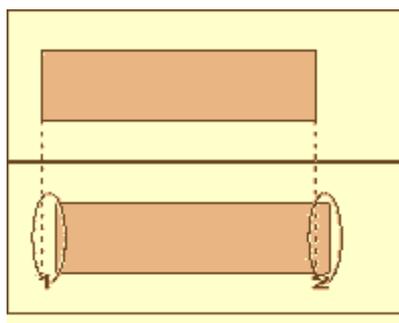


Рис.2

Конечно, в фильме существует много ситуаций, связанных со сменой действия, когда первый кадр новой сцены настолько отличается от предыдущего, что его проще сделать ключевым, чем разностным. Может показаться, что в компьютерном фильме будет столько ключевых кадров, сколько новых ракурсов камеры. Тем не менее, их гораздо больше. Регулярное расположение подобных кадров в потоке позволяет пользователю оперативно начинать просмотр с любого места фильма: "если пользователь решил начать просмотр фильма с середины, вряд ли он захочет ждать, пока программа распаковки вычислит все разности с самого начала" Кроме того, указанная профилактическая мера позволяет эффективно восстановить изображение при любых сбоях или при "потере темпа" и пропуске отдельных кадров на медленных компьютерных системах.

Заметим, что в современных методах сохранения движущихся видеоизображений используются и другие типы кадров.

Существует множество различных форматов представления видеоданных. В среде Windows, например, уже более 10 лет (начиная с версии 3.1) применяется формат Video for Windows, базирующийся на универсальных файлах с расширением AVI (**Audio Video Interleave - чередование аудио и видео**). Суть AVI файлов состоит в хранении структур произвольных мультимедийных данных, каждая из которых имеет простой вид,

изображенный на рис.3. Файл как таковой представляет собой единый блок, причем в него, как и в любой другой, могут быть вложены новые блоки. Заметим, что идентификатор блока определяет тип информации, которая хранится в блоке.

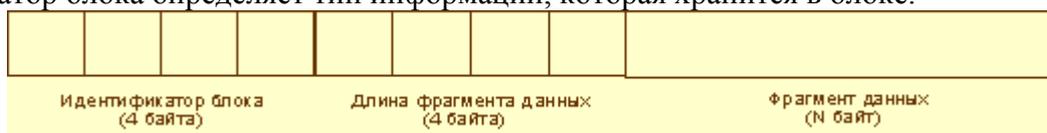


Рис.3

Внутри описанного выше своеобразного контейнера информации (блока) могут храниться абсолютно произвольные данные, в том числе, например, блоки, сжатые разными методами. Таким образом, все AVI-файлы только внешне выглядят одинаково, а внутри могут различаться очень существенно.

Еще более универсальным является мультимедийный формат **Quick Time**, первоначально возникший на компьютерах Apple. По сравнению с описанным выше, он позволяет хранить независимые фрагменты данных, причем даже не имеющие общей временной синхронизации, как этого требует AVI. В результате в одном файле может, например, храниться песня, текст с ее словами, нотная запись в MIDI-формате, способная управлять синтезатором, и т.п. Мощной особенностью Quick Time является возможность формировать изображение на новой дорожке путем ссылок на кадры, имеющиеся на других дорожках. Полученная таким способом дорожка оказывается несоизмеримо меньше, чем если бы на нее были скопированы требуемые кадры. Благодаря описанной возможности файл подобного типа легко может содержать не только полную высококачественную версию видеofilма, но и специальным образом "упрощенную" копию для медленных компьютеров, а также рекламный ролик, представляющий собой "выжимку" из полной версии. И все это без особого увеличения объема по сравнению с полной копией.

Все большее распространение в последнее время получают системы сжатия видеоизображений, допускающие некоторые незаметные для глаза искажения изображения с целью повышения степени сжатия. Наиболее известным стандартом подобного класса служит **MPEG (Motion Picture Expert Group)**, который разработан и постоянно развивается созданным в 1988 году Комитетом (группой экспертов) международной организации ISO/IEC (International Standards Organization/International Electrotechnical Commission) по стандартам высококачественного сжатия движущихся изображений. Методы, применяемые в MPEG, непросты для понимания и опираются на достаточно сложную математику. Укажем лишь наиболее общие приемы, за счет которых достигается сжатие. Прежде всего, обрабатываемый сигнал из RGB-представления с равноправными компонентами преобразуется в яркость и две "координаты" цветности. Как показывают эксперименты, цветовые компоненты менее важны для восприятия и их можно проредить вдвое. Кроме того, производится специальные математические преобразования (DCT - дискретно-косинусное преобразование), несколько загроубляющее изображение в мелких деталях. Опять таки из экспериментов следует, что на субъективном восприятии изображение это практически не сказывается. Наконец, специальными методами (в том числе и методом, изображенным на рис.2) ликвидируется сильная избыточность информации, связанная со слабыми отличиями между соседними кадрами. Полученные в результате всех описанных процедур данные дополнительно сжимаются общепринятыми методами, подобно тому, как это делается при архивации файлов.

В последнее время все большее распространение получает технология под названием **DivX** (происходит от сокращения слов **Digital Video Express**, обозначающих название видеосистемы, которая "прославилась" неудачной попыткой взимать небольшую плату за каждый просмотр видеодиска; к собственно технологии DivX это никакого отношения не имело). Благодаря DivX удалось достигнуть степени сжатия, позволившей вместить качественную запись полнометражного фильма на один компакт-диск - сжать 4,7 Гб DVD-фильма до 650 Мб. И хотя это достижение, к сожалению, чаще всего используется для пиратского копирования, сам по себе этот факт не умаляет достоинств новой технологии. Как и то, что самая первая версия сжатия DivX была сработана французскими хакерами из MPEG-4 - современные версии DivX уже не имеют к этому событию никакого отношения.

Наиболее популярные программы проигрывания видеофайлов позволяют использовать замещаемые подсистемы сжатия и восстановления видеоданных - **кодеки (от англ. compression/decompression - codec**, сравните с образованием термина "модем"). Такой подход позволяет легко адаптировать новые технологии, как только те становятся доступными. Замещаемые кодеки хороши как для пользователей, так и для разработчиков программного обеспечения. Тем не менее, большое разнообразие кодеков создает определенные трудности для производителей видеопродукции. Часто в качестве выхода из создавшегося положения необходимые кодеки помещают на компакт-диск с фильмами или даже поставляют видеоматериалы в нескольких вариантах, предоставляя тем самым возможность выбрать подходящий. Все больше распространяется автоматизация распознавания, когда плеер, обнаружив информацию об отсутствующем кодеке, загружает его из Интернет.