

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ**

**ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО**

**ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ**

**«Донской государственный технический университет»**

**(ДГТУ)**

**Доклад на тему**

**«Современные алгоритмы сжатия видео»**

(теория информации и кодирования)

Выполнил: студент ВМО21  
Оганесьянц К.П.

Проверил: Ляхницкая О.В.

Ростов-на-Дону

2021

**Введение**

Количество различных форматов представления цифрового видео в своем многообразии может остановить рядового пользователя в выборе определенного и подходящего для него формата. Задача этого доклада – провести обзор и сравнительную характеристику существующих алгоритмов сжатия цифровой видео-информации

**Основная часть**

**AVI (Audio Video Interlive)**

AVI-файлы — особый случай файлов RIFF. (сокращенно от Resource Interchange File Format). Этот формат, изначально предназначавшийся для обмена мультимедийными данными, был создан Microsoft совместно с IBM. Данный формат является наиболее распространенной формой представления видео на персональных компьютерах. В зависимости от формы представления видеоданных файлы AVI бывают различных стандартов.

|  |  |
| --- | --- |
|  |  |

***Стандарты AVI-файлов***

Видеоданные в AVI могут быть сформированы и сжаты различными способами. Например, Video for Windows 1.1поставляется со стандартным набором компрессоров. А именно:

Intel Indeo (version 3.2);

Microsoft Video 1;

Microsoft RLE (Run Length Encoding);

Cinepak;

Для каждого стандарта предусмотрен так называемый 4-символьный код The Four Character Code (FOURCC).

Полные кадры (несжатые)

Можно держать AVI-файлы в форме несжатых кадров. Для этого не нужно никаких компрессоров.

4-символьный код этого стандарта — ‘DIB’, (Microsoft Device Independent Bitmap).

Тем не менее существует по крайней мере еще три FOURCC для несжатых AVI-файлов:

'RGB '; 'RAW '; 0x00000000 (FOURCC, чье шестнадцатеричное значение равно 0)

Кодирование переменной длины (MRLE)

Особенностью данной схемы кодирования видео изображений является то, что MRLE способна кодировать видео только в 8-битной палитре (256 цветов). Она не поддерживает режимы «High Color» (2^16 цветов) или «True Color» (2^24 или 2^32 цветов).

Microsoft Video 1 (MSVC или CRAM)

Microsoft Video 1 поддерживает только 8- или 16-битную цветовую палитру.

Intel Real Time Video 2.1 (Indeo 2.1) (RT21)

Включен в состав стандартной поставки Microsoft Video For Windows 1.1.

Intel Indeo 3.1/3.2 (IV31 и IV32)

Intel® Indeo® Media Kit 5.0

Cinepak (CVID)

Один из наиболее распространенных и используемых компрессоров Video for Windows. Обеспечивает наиболее быстрое воспроизведение видео. В отличие от Indeo 32, которая обеспечивает чуть лучшее качество, однако заметно отягощает процессор при декомпрессии, Cinepak максимально разгружает процессорный ресурс.

На сегодняшний день существует по крайней мере три стандарта для Cinepak.

Cinepak SuperMac (изначальный, 16-битный компрессор);

Cinepak Radius (новый, улучшенный 16-битный компрессор);

Cinepak Radius[32] (32-битная версия Radius Cinepak, поставляемая совместно с Windows 95).

Особенностью метода является использование алгоритма векторной квантизации изображений совместно с алгоритмом разницы кадров.

Один из наиболее популярных форматов представления цифрового видео. В настоящее время существуют три его спецификации: MPEG-1, MPEG-2 и MPEG-4. Несмотря на большую гибкость стандарта, позволяющую для различных приложений менять в широкий пределах значения большинства его параметров (такие как разрешение изображений, аспектное отношение, частота кадров), разработчики его первоначально были ориентированы на использование в качестве основного носителя кодированной информации компакт дисков (CD-ROM) со скоростью передачи данных 50 Кбит/с. В результате базовый алгоритм (MPEG 1) ограничивает скорость передачи диапазоном 150-225 Кбит/с, разрешение изображений (кадров) как 352\*288(Pal) или 320\*240 (NTSC), частоту их смены 25 (PAL) или 30 (NTSC). Далее для простоты ограничимся рассмотрением Pal-системы.

*MPEG-компрессия использует следующие основные идеи:*

Устранение временной избыточности видео, учитывающее тот факт, что в пределах коротких интервалов времени большинство фрагментов сцены оказываются неподвижными или незначительно смещаются по полю.

Устранение пространственной избыточности изображений подавлением мелких деталей сцены, несущественных для визуального восприятия человеком;

Использование более низкого цветового разрешения при yuv-представлении изображений (y — яркость, u и v — цветоразностные сигналы) — установлено, что глаз менее чувствителен к пространственным изменениям оттенков цвета по сравнению с изменениями яркости.

Повышение информационной плотности результирующего цифрового потока путем выбора оптимального математического кода для его описания (например, использование более коротких кодовых слов для наиболее часто повторяемых значений).

Изображения в Mpeg-последовательности подразделяются на следующие типы:

I (intra), играющие роль опорных при восстановлении остальных изображений по их разностям;

P (predicted), содержащие разность текущего изображения с предыдущим I или P с учетом смещений отдельных фрагментов;

B (bidirectionally predicted), содержащие разность текущего изображения с предыдущим и последующим изображениями типов I или P с учетом смещений отдельных фрагментов.

Изображения объединяются в группы (GOP — Group Of Pictures), представляющие собой минимальный повторяемый набор последовательных изображений, которые могут быть декодированы независимо от других изображений в последовательности. Типичной является группа вида (I0 B1 B2 P3 B4 B5 P6 B7 B8 P9 B10 B11) (I12 B13 B14 P15 B16 B17 P18…), в которой I тип повторяется каждые полсекунды. Обратим внимание, что в изображении P3 основная часть фрагментов сцены предсказывается на основании соответствующих смещенных фрагментов изображения I0. Собственно кодированию подвергаются только разности этих пар фрагментов. Аналогично P6 «строится» на базе P3, P9 — на базе P6 и т.д. В то же время большинство фрагментов B1 и B2 предсказываются как полусумма смещенных фрагментов из I0 и P3, B4 и B5 — из P3 и P6, B7 и B8 — из P6 и P9 и т.д. Наряду с этим B-изображения не используются для предсказания никаких других изображений. В силу зависимости изображений в процессе их кодирования меняется порядок следования. Для вышеприведенной последовательности он будет следующим: I0 P3 B1 B2 P6 B4 B5 P9 B7 B8 I12 B10 B11 P15 B13 B14 P18 B16 B17…

Ясно, что точность кодирования должна быть максимальной для I, ниже — для P, минимальной — для B. Установлено, что для типичных сцен хорошие результаты достигаются при отведении числа бит для I в 3 раза больше, чем для P , и для P в 2-5 раз больше, чем для B. Эти отношения уменьшаются для динамичных сцен и увеличиваются для статичных.

Отдельные изображения состоят из макроблоков.

**Макроблок** — это основная структурная единица фрагментации изображения. Он соответствует участку изображения размером 16\*16 пикселов. Именно для них определяются вектора смещения относительно I- или P-изображений. Общее число макроблоков в изображении — 396. Для повышения устойчивости процесса восстановления изображений к возможным ошибкам передачи данных последовательные макроблоки объединяют в независимые друг от друга разделы (slices), максимальным числом 396. В предельном случае «чистой» передачи на изображение приходится всего один раздел из 396 макроблоков. В свою очередь каждый макроблок состоит из шести блоков, четыре из которых несут информацию о яркости Y, а по одному определяют цветовые U- и V-компоненты. Каждый блок представляет собой матрицу 8\*8 элементов. Блоки являются базовыми структурными единицами, над которыми осуществляются основные операции кодирования, в том числе выполняется дискретное косинусное преобразование (DCT — Discrete Cosine Transform) и квантование полученных коэффициентов.

В заключение подчеркнем важную особенность MPEG-последовательностей - они не допускают покадрового редактирования. Действительно, в силу вышеописанной сложной зависимости изображений внесение изменений в одно из них неизбежно повлечет за собой необходимость комплексного пересчета множества параметров во всей группе.

**Motion JPEG или M-JPEG (MJPG)**

Большинство настольных систем видеозаписи (video capture) и видеоредактирования (video editing) используют именно этот метод при записи видео в AVI-файлы. В Motion JPEG каждый видеокадр изображения сжимается отдельно с использованием стандарта JPEG. Никаких других дополнительных алгоритмов при этом не используется. Безусловным достоинством этого метода является возможность редактирования видео без потерь качества, так как кадры являются независимыми. Этим, по сути дела, и определяется использование данного метода именно как механизма хранения видео, служащего для его редактирования, а не для распространения. Motion JPEG использует алгоритм блокового ДКП (Block Discrete Cosine Transform (DCT)) для сжатия изображений.

**Editable MPEG (XMPG. MPGI)**

Editable MPEG, так же как и M-JPEG, используется для редактирования цифрового видео представляет собой AVI-файл, состоящий только из кадров MPEG типа i. Однако все другие механизмы сжатия MPEG тут задействованы. Входит в стандартную поставку Microsoft Video for Windows 1.1. и используется такой настольной системой редактирования цифрового видео, как, например, Adobe Premiere.

Хотя вышеперечисленные компрессоры достаточно популярны, тем не менее это далеко не все стандарты сжатия AVI-файлов. Характеризуя эту группу компрессоров, можно отметить, что они проектировались и создавались в первую очередь как средства сжатия видео- и аудиоданных, хранящихся на жестких дисках и компакт-дисках, а это, в свою очередь, свидетельствует об их небольших возможностях при сжатии и относительно высоком качестве при воспроизведении.

С приходом Интернета все большую популярность получают методы и средства сжатия видео- и аудиоданных, позволяющих, применяя передовые технологии (sophisticated motion estimation and compensation, wavelets, fractals и другие), достичь наименьших соотношений «килобит/секунда», позволивших проводить, например, сеансы видеоконференций средствами Интернета. Ясно, что такие методы сжатия обеспечивают существенно большую степень сжатия, при относительно низком качестве.

**Поколение компрессоров — H.XXX**

Совсем недавно (18/5/97) была официально зарегистрирована новая серия компрессоров цифрового видео, определяющая тенденции развития механизмов сжатия видео. Некоторые компрессоры семейства H.XXX, например H.261, довольно популярны, другие малоизвестны и используют такие передовые и улучшенные технологии, как wavelets.

Отличительной особенностью компрессоров семейства H.XXX является их нацеленность на уменьшение потока цифрового видео через Интернет, что естественно приводит к отходу фактора качества на второй план. Некоторые продукты семейства H.XXX Microsoft совершенствует, а некоторые уже входят в состав таких пакетов телеконференций, как NetShow и NetMeeting.

**VDOWave или VDOLive от VDONet (VDOM, VDOW)**

VDONet выпускает wavelet-основанный видеокомпрессор, включенный в комплект реализации 32-битной версии Video for Windows. Microsoft использует VDOWave как часть NetShow. В настоящее время существуют две версии компрессора VDOWave:

VDOWave 2.0 is a fixed rate video codec.

VDOWave 3.0 is a "scalable" video codec.

Стандартный набор NetShow 2.0 устанавливает только декодер VDOWave. Средство разработки NetShow 2.0 устанавливает как кодер, так и декодер VDONet VDOWave. По некоторым тестам, VDOWave существенно превосходит по сжатию MPEG-1 и другие компрессоры, базирующиеся на алгоритме блокового ДКП (block Discrete Cosine Transform), но лишь при низких отношениях «килобит/секунда».

**MPEG-4 (MPG4)**

Microsoft's NetShow 2.0 устанавливает компрессор MPEG-4, являющийся новым международным стандартом, который, однако, официально еще не признан.

Технологии сжатия цифрового видео

Существует множество технологий сжатия цифрового видео, однако остановимся на тех из них, которые легли в основу наиболее популярных компрессоров видео. Некоторые из рассматриваемых компрессоров используют не одну технологию сжатия, а некоторую их совокупность. Например, и Indeo 3.2, и Cinepak используют векторную квантизацию. Международные стандарты MPEG-1, MPEG-2, MPEG-4, H.261 и H.263 используют комбинированную технологию БДКП и компенсацию движения. Некоторые современные алгоритмы используют технологию ДВП (Discrete Wavelet Transform, или DWT). Другие технологии включают Фрактальное сжатие изображений (Fractal Image Compression).

Для начала давайте определим понятие «качество сжатия видео».

**Сжатие без потерь качества**

(1) Сжатие изображений может осуществляться без потерь качества лишь в том случае, если в процессе сжатия не было потерь данных. В результате полученное после декомпрессии изображение будет в точности (побитно) совпадать с оригиналом. Примером такого сжатия может служить формат GIF для статической графики и GIF89a для видео.

**Сжатие с потерями качества**

(2) Сжатие может происходить с потерями качества, если в процессе сжатия информация была потеряна. Однако с точки зрения человеческого восприятия сжатием с потерями следует считать лишь такое сжатие, при котором возможно на глаз отличить результат сжатия от оригинала. Таким образом, несмотря на то что два изображения — оригинал и результат сжатия с использованием того или иного компрессора — побитно могут не совпадать, тем не менее разница между ними может быть совсем незаметной. Примером может служить алгоритм JPEG для сжатия статической графики и алгоритм M-JPEG для сжатия видео.

**Сжатие без потерь с точки зрения восприятия**

(3) Формально являясь сжатием с потерями качества, схема сжатия может в то же время казаться сжатием без потерь с точки зрения восприятия ее человеком. Большинство технологий сжатия с формальной потерей качества имеют так называемый Фактор Качества Сжатия (ФКС), характеризующий именно воспринимаемую сторону качества и варьирующийся в пределах от 0 до 100. При факторе качества сжатия равном 100 воспринимаемые характеристики качества сжатого видео неотличимы от оригинала.

**Сжатие с естественной потерей качества**

(4) JPEG и MPEG и другие технологии сжатия с потерей качества иногда сжимают, без потерь переступая за грань сжатия с точки зрения восприятия видеоинформации. Тем не менее сжатые видео и статические изображения вполне приемлемы для адекватного восприятия их человеком. Иными словами, в данном случае наблюдается так называемая естественная деградация изображения, при которой теряются некоторые мелкие детали сцены. Похожее может происходить и в естественных условиях, например при дожде или тумане. Изображение в таких условиях, как правило, различимо, однако детализация его уменьшается.

**Сжатие с неестественными потерями качества**

(5) Низкое качество сжатия, в значительной степени искажающее изображение и вносящее в него искусственные (не существующие в оригинале) детали сцены, называется неестественным сжатием с потерей качества. Примером тому может служить некоторая «блочность» в сильно сжатом MPEG-е и в других компрессорах, использующих технологию БДКП. Неестественность заключается в первую очередь в нарушении самых важных с точки зрения восприятия человеком характеристик изображения — контуров. Опыт показывает, что именно контуры позволяют воспринимающему аппарату человека правильно идентифицировать тот или иной визуальный объект.

**Обзор технологий и алгоритмов сжатия видео**

**Run Length Encoding**

Компрессорами, использующими технологию RLE, являются:

**Microsoft RLE** (MRLE) RLE используется также для кодирования коэффициентов в БДКП, применяющемся в MPEG-1234, H.261, H.263 и JPEG.

Достоинства и недостатки

Работает исключительно с 8-битовыми изображениями.

Не подходит для сжатия полноцветного видео.

Обзор

RLE кодирует последовательность повторяющихся элементов изображения или одноцветных элементов одним кодовым словом. Например, последовательность элементов изображения 77 77 77 77 77 77 77 может быть закодирована как 7 77 (для семи 77-рок). RLE хорошо сжимает изображения, в которых наблюдается повторение контуров или цветов отдельных элементов. В полноцветных изображениях повторений цвета значительно меньше, поэтому сжатие полноцветного видео с использованием технологии RLE лишено всякого смысла.

**Векторная квантизация (Vector Quantization,VQ)**

Компрессорами, использующими технологию VQ, являются Indeo 3.2 и Cinepak. Оба они применяют цветовую схему YUV (а не RGB).

Достоинства и недостатки

Процесс кодирования очень трудоемок и практически неосуществим без специального дополнительного оборудования.

Процесс декодирования очень быстр.

Блоковые искажения при высоких коэффициентах сжатия.

Технологии, использующие алгоритмы БДКП, ДВП могут достигать более высоких уровней сжатия.

Обзор

Основная идея векторной квантизации заключается в разбиении изображения на блоки (размером 4x4 пиксела в цветовой схеме YUV для компрессоров Indeo и Cinepak). Как правило, некоторые блоки оказываются похожими друг на друга. В этом случае компрессор идентифицирует класс похожих блоков и заменяет их одним общим блоком. Кроме того, генерируется двоичная таблица (карта) таких общих блоков из самых коротких кодовых слов. VQ-декодер затем, используя таблицу, собирает изображение поблочно из общих блоков. Ясно, что данный способ кодирования с потерями качества, так как, строго говоря, схожесть блоков весьма относительна. Здесь допускается аппроксимация реальных блоков изображения к общему, их объединяющему. Процесс кодирования длителен и трудоемок, так как кодеру необходимо выявлять принадлежность каждого блока изображения к какому-нибудь общему блоку. Однако задача декодирования в этом случае сводится к задаче построения изображения по заданной карте из общих блоков и не занимает много аппаратных и временных ресурсов. Таблицу или карту также называют еще и кодовой книгой, А двоичные коды, входящие в нее, — кодовыми словами, соответственно. Наибольшее сжатие с использованием алгоритма VQ достигается путем уменьшения числа классов общих блоков, то есть предположением о схожести относительно большего числа блоков изображения, и, как следствие, уменьшением кодовой книги. По мере уменьшения размеров кодовой книги качество воспроизводимого видео ухудшается. В результате на изображении появляется искусственная «блочность». (Vector Quantization and Signal Compression, A. Gersho and R. Gray, Boston, MA : Kluwer, 1992).

**Дискретное Косинусное Преобразование (ДКП)**

Компрессоры, использующие ДКП: Motion JPEG; Editable MPEG; MPEG-1; MPEG-2; MPEG-4.

Достоинства и недостатки

«Блочность» при высокой компрессии.

Закругление острых углов изображения. Случайное «размывание» острых краев изображений.

Кодирование очень трудоемко. Только в последнее время удалось осуществить процесс кодирования программно, а не аппаратно.

Обзор

ДКП является широко используемым при сжатии изображений преобразованием. Стандарт сжатия статической графики JPEG, используемый в видеоконференциях стандарт H.263, цифровые видеостандарты MPEG (MPEG-1, MPEG-2 и MPEG-4) — все они используют ДКП. В этих стандартах используется, в частности, 2-мерное ДКП, применяемое последовательно к блокам изображения размерностью 8 x 8 пикселов. ДКП вычисляет 64 (8x8 = 64) коэффициента, которые затем квантизуются, обеспечивая тем самым реально сжатие. В большинстве изображений большинство ДКП-коэффициентов в силу своей малости после квантизации обнуляется. Это свойство ДКП и лежит в основе множества алгоритмов сжатия, использующих ДКП.

Вдобавок известно, что человеческий глаз гораздо менее чувствителен к высокочастотным компонентам изображения, представляемым большими коэффициентами ДКП. К этим большим значениям коэффициентов может быть применен (и, как правило, применяется) больший фактор квантизации. В частности, матрица 64 факторов квантизации для каждого из 64 коэффициентов ДКП, применяемая в алгоритме JPEG, имеет большие факторы квантизации для коэффициентов ДКП, соответственно, большей частоты. После квантизации коэффициенты подвергаются алгоритму RLE. Далее для частых комбинаций используются короткие кодовые слова, для более редких — относительно длинные. Осуществляется вероятностное кодирование.

ДКП, в свою очередь, лучше всего объяснять на примере одномерного ДКП. Двухмерное ДКП представляет собой одномерное ДКП, применяемое последовательно для каждого ряда (строки) блока пикселов и каждой колонки блока пикселов, полученного от одномерного ДКП строк. Одномерное ДКП, применяемое к N выборкам (пикселам в изображении или выборкам в звуковом файле). ДКП есть матрица размерности NxN, строки которой представляют собой косинусные функции:

ДКП(m,n) = sqrt( (1 - delta(m,1) ) / N ) \* cos( (pi/N) \* (n - 1/2) \* (m-1) )

, где

ДКП (m,n) есть одномерная матрица ДКП

m, n = 1,...,N

pi = 3.14159267...

N = число выборок в блоке

delta(m,1) = 1 если m = 1 и 0 в противном случае

cos(x) = косинус x, измеряемый в радианах.

**Дискретное Wavelet-преобразование (DWT)**

Компрессоры, использующие DWT (Discrete Wavelet Transform): Intel Indeo 5.x; Intel Indeo 4.x

Достоинства и недостатки

Большинство как статических, так и динамических изображений, сжатых при помощи алгоритма DWT, не имеет характерной для алгоритма ДКП блочной структуры.

Относительное качество изображений, сжатых с использованием DWT, превосходит качество изображений, сжатых при помощи ДКП, при тех же коэффициентах сжатия.

DWT несколько размазывает, закругляет острые контуры изображения. Так называемый контурный шум или эффект Гиббса.

Обзор

DWT-алгоритм основан на передаче сигнала, например изображения, через пару фильтров: низкочастотный и высокочастотный. Низкочастотный фильтр выдает грубую форму исходного сигнала. Высокочастотный фильтр выдает сигнал разности или дополнительной детализации.

В свою очередь, результат на выходе высокочастотного фильтра (добавочный сигнал детализации) может быть подвернут той же процедуре и так далее.

Выход низкочастотного фильтра, как уже отмечалось, представляет собой грубую аналогию исходного сигнала. Если исходным сигналом является изображение, то на выходе низкочастотного фильтра получится расплывчатое, размытое изображение с низким разрешением. Выход высокочастотного сигнала добавляет детали к изображению. В сочетании с выходом низкочастотного фильтра может быть воспроизведено, таким образом, исходное изображение. Грубая форма исходного сигнала (сигнал на выходе низкочастотного фильтра) иногда называют основным уровнем (base layer), а дополнительный сигнал детализации — уровнем улучшения (enhancement layer). Сигнал на выходе высокочастотного фильтра h[n] может быть пропущен снова через пару фильтров, и процесс, таким образом, может быть повторен, пока не будет достигнута достаточная степень детализации исходного сигнала x[n]. Однако ясно, что никакого сжатия здесь не достигается. Преобразование попросту воспроизводит то же количество битов, которое было в исходном сигнале. Выходные значения называются коэффициентами преобразования, или коэффициентами wavelet-преобразования.

Преобразование Хара используется в основном в области сжатия изображений. Для других целей используются более сложные фильтры преобразований. Сжатие же достигается в основном за счет применения некоторой формы квантизации (скалярной или векторной) к добавочному сигналу детализации. Далее к полученным коэффициентам преобразования применяется техника вероятностного (энтропийного) кодирования.

Допустим, что в приведенном выше примере входной сигнал x[n] представляет собой последовательность 8-битных выборок растра полутонового изображения. Для выхода низкочастотного фильтра g[n] теперь можно использовать те же 8 бит, а для высокочастотного h[n] — уже меньше, например 4. Это, по сути, скалярная квантизация. Далее выход высокочастотного фильтра будет стремиться к нулю, так как коэффициенты преобразования будут убывать по мере применения алгоритма. Таким образом, возможно применить вероятностное кодирование к сигналу детализации h[n].

**Разница кадров**

Компрессорами, использующими технологию разницы кадров, являются: Cinepak

Достоинства и недостатки

В целом может обеспечивать сжатие, лучшее, чем независимое сжатие отдельных кадров.

Возникающие в ходе кодирования ошибки накапливаясь, требуют наличия дополнительного ключевого кадра.

Обзор

Алгоритм разницы кадров использует то обстоятельство, что во многих видео изображение от кадра к кадру мало чем различается. По мере применения алгоритма векторной квантизации для кодирования каждого следующего кадра и получения при этом малых коэффициентов, которые трудно кодируются, в кадры постепенно вкрадывается ошибка. Это требует включения в видеоряд так называемых ключевых кадров, которые кодируются без учета предыдущих и являются так называемыми «опорными точками» в видео.

**Компенсация движения**

Компрессорами, использующими технологию компенсации движения, являются: MPEG-1,2 и 4.

Достоинства и недостатки

По сравнению с механизмом разницы кадров механизм компенсации движения позволяет достигать большей степени сжатия.

Кодирование весьма трудоемко и требует специальной аппаратуры.

Технология компенсации движения используется в таких международных стандартах сжатия цифрового видео, как: MPEG, H.261 и H.263.

Наибольшее сжатие достигается в сценах с пониженным движением.

Обзор

Компенсация движения основана на использовании ряда сложных алгоритмов. Сфера, где данная технология сжатия эффективна, как правило, сводится к видеоряду, в котором объект изменяет свое местоположение относительно неподвижного фона. Объекты, изменяющиеся по форме, приближающиеся или удаляющиеся (движущаяся камера), не подлежат эффективному сжатию посредством алгоритма компенсации движения. Сжатие возможно заданием вектора смешения элементов изображения вместо хранения больших значений новых координат данных элементов изображения. Основным блоком (относительно которого задается вектор смещения остальных блоков) может являться любой блок изображения размером 16x16 пикселов, максимально похожий на кодируемый (предсказываемый) блок. Ясно, что кадр, на который ссылаются таким образом другие кадры, должен быть декодирован ранее. Однако совсем не обязательно, чтобы опорный кадр предшествовал предсказываемому кадру. MPEG позволяет производить предсказание в обоих направлениях путем введения так называемых B- (bi-directionally predicted) кадров.