

**Московский авиационный институт
(государственный ТѢХнический университет)**

Факультет прикладной математики

Кафедра вычислительной математики и программирования

**Лекции по курсу «Методы и Средства
Мультимедиа»**

Преподаватель: А. В. Крапивенко
Студент: И. К. Никитин

Москва, 2010

Содержание

1	Ресурсоемкость видео	5
2	Форматы видео	5
3	Формат AVI	6
3.1	Структура файла	6
3.2	Работа Windows Media Player	6
3.3	Кодеки	6
3.4	Чего нет в AVI	6
4	Формат QuickTime	7
4.1	Что есть	7
5	Формат телевидения	7
6	Говорящая голова	8
6.1	Межкадровое кодирование	8
6.2	Векторное кодирование	8
7	Требования приложений к алгоритмам сжатия	8
8	Критерии сравнения	8
9	Монтажные эффекты и видео переходы	9
9.1	Таимкоды	9
9.2	Ленты	9
9.3	Способы монтажа видео переходов	9
9.4	Видео монтажные приемы	9
9.5	Виртуальные студии	9
9.6	Виртуальные актеры	9
10	Алфавитные методы	11
10.1	Алгоритм построения кодов Хаффмана	13
10.2	Коды Хаффмана с фиксированный таблицей	14
11	Алфавитные методы	15
11.1	Коды грея посмотреть	16
12	Алгоритмы вероятностного кодирования	16

13 Физиологические предпосылки для методов сжатия с частичной потерей качества оригинала	16
14 Законы субъективного восприятия яркости	16
14.1 Закон Блоха-Шарпантье	16
15 Метод оценки качества	17
15.1 Субъективные методы оценки качества	17
15.1.1 Сложности тестирование	17
15.1.2 Шкала деградации с двумя стимулами	18
15.1.3 Непрерывная оценка качества с одним стимулом.	18
15.1.4 Попарное сравнение	18
15.2 Дополнительные размерности оценок	18
15.3 Объективные методы	19
15.4 Объективные методы	20
15.4.1 Психофизиологическая кривая	20
16 Контекстные методы	20
16.1 JND	20
16.2 PDM	21
16.3 SSIM	21
17 По курсовой работе	21
17.1 Реферат	21
17.2 Вычислительная часть	21
17.3 Аналитическая часть	21
Методы сжатия с потерями	23
18 VQ - Метод векторной квантизации	23
19 Временное квантование	23
20 Контурное кодирование	23
21 Цветовое прореживание	23
21.1 JPEG	26
Группа MPEG-методов	27
22 MPEG-1	27

23 MPEG-2	27
24 Временная избыточность	28
25 Компенсация движения	29
25.1 Метод прямого сравнении блоков изображении	29
26 Методы субоптимального поиска	30
26.1 Сигнатурные методы	30
26.2 Поиск с разбавленным расстоянием	30
26.3 Локализованный телескопический поиск	31
26.4 Двумерный логарифмический поиск	32
26.5 Многомерный поиск	34
26.6 Метод корреляции фаз	35
27 Характерные дефекты MPEG	36
28 MPEG-4	36
28.1 MPEG-4	36
Метод кросс-кадровой интерполяции	37
29 Статистика распределения квантов по классам	38
30 Преимущества и недостатки	38
Фрактальное сжатие	40
31 Построение алгоритма	40
Мат аппарат	43
32 Th1	43
33 Th2	43
34 Th3 О коллаже	43
Фрактальное сжатие видео	45

Вейвлеты (wavelets)	46
Вейвлеты (wavelets)	47
35 Базис Хаара	49
Eye-Tracker — системы отслеживания движений зрачков глаз	51
Интернет вещание	52
36 Файловое	52
37 Потокное	52
38 Подготовка к вещанию	52
39 Виды соединений с клиентом	52
Телевидение	55
40 форма телевизионного экрана	55
41 National Television Standards Committee	55
42 phase-alternating line (PAL)	56
43 SECAM	56

1. Ресурсоемкость видео

Цифровое видео крайне ресурсоемко. Это не сложно понять если попытаться рассчитать сколько по объему (в гигабайтах) занимает один час телевизионного вещания.

Размер кадра для телевизионного вещания составляет 720×576 . Так же учтем, что картинка имеет три составляющих (об этом будет рассказано ниже, сразу оговоримся, что эти составляющие, не RGB, столь привычные современному человеку). Вещание ведется с частотой 25 кадров в секунду.

Перемножая все эти параметры друг на друга и на время, можно получить что один час телевизионного вещания составляет 104 Гб.

$$\begin{aligned} 720 \times 576 \cdot 3 \cdot 25 \cdot 3600 &= \\ &= 111974400000 \text{ б} \approx 104 \text{ Гб} \end{aligned}$$

Такая цифра даже по сегодняшним меркам является недопустимой для хранения видео-данных. Отсюда возникает необходимость сжатия.

2. Форматы видео

Перечислим существующие на сегодняшний день форматы видео-файлов:

*. **3gp** — мультимедийный контейнер, определяемый Партнёрским Проектом Третьего поколения¹ для мультимедийных служб 3G UMTS. Многие современные мобильные телефоны имеют функции записи и просмотра аудио и видео в этом формате.

*. **flv** — Flash Video — медиаконтейнер, используемый для передачи видео через Интернет. Формат используется в основном для передачи фотореалистичных изображений.

*. **flc** — Flash Cartoon — медиаконтейнер, используемый для передачи видео через Интернет. Формат используется в основном для передачи мультипликационных изображений.

*. **avi** — Audio Video Interleave (сокращённо AVI; букв. «чередование аудио и видео»).

*. **mkv** — Matroska (Matröška, Матрёшка) — проект, нацеленный на создание открытого, гибкого, кроссплатформенного (включая аппаратные платформы) формата мультимедийного контейнера и набора инструментов и библиотек для работы с данными в этом формате. Этот проект является развитием проекта MCF, но значительно отличается тем, что основан на EBML (Extensible Binary Meta Language — расширяемый двоичный метаязык) — двоичном аналоге языка XML. Использование EBML позволяет расширять формат без потери совместимости со старыми программами.

*. **mov** — QuickTime — проприетарная технология компании Apple, разработанная в 1991 году для воспроизведения цифрового видео, звука, текста, анимации, музыки и панорамных изображений в различных форматах.

*. **mp*** — ???

*. **ogg** — открытый стандарт формата мультимедиа-контейнера, являющийся основным файловым и потоковым форматом для мультимедиа-кодексов фонда Xiph.Org. Как и все технологии, разрабатываемые под эгидой Xiph.Org, он является открытым и свободным стандартом, не имеющим патентных или лицензионных ограничений.

*. **rm** — ???

*. **vob** — формат файлов, используемый для хранения DVD-Video. VOB - это мультимедийный контейнерный формат файла, основанный на MPEG-2, и способный содержать в себе несколько потоков видео/аудио, субтитры, а также меню фильма.

*. **wmv** — Windows Media Video — название системы видео кодирования, разработанной компанией Microsoft для хранения и трансляции видеoinформации в проприетарных форматах Microsoft. Использует расширение имени файла .wmv. Входит в мультимедийный пакет Windows Media.

¹Third Generation Partnership Project (3GPP)

Если рассматривать приведенные форматы с точки зрения сжатия то можно сказать следующее:

.mp — используют дискретное косинусное преобразование. Это алгоритм с потерей качества.

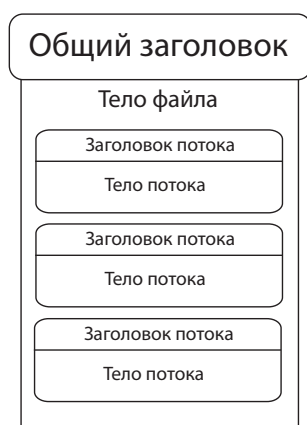
*.flv и *.flc — межкадровое кодирование без потери качества.

*.wmv — используются вейвлеты.

3. Формат AVI

Audio Video Interleave — RIFF-медиаконтейнер, впервые использованный Microsoft в 1992 году. RIFF² — один из форматов файлов-контейнеров для хранения потоковых мультимедиа-данных (видео, аудио, возможно текст). Наиболее известными форматами, использующими RIFF в качестве контейнера, являются: AVI (видео), WAV (аудио), RMI (MIDI-треки).

3.1. Структура файла



Файл делится на заголовок и тело. Тело файла может содержать до трех потоков, которые в свою очередь тоже имеют заголовок и тело.

В общем заголовке содержатся:

- 1) Количество потоков.
- 2) Абсолютные адреса начала потоков.

В заголовках потоков в числе прочей информации содержится FOURCC³ — четырехсимвольный код кодека. Внутри каждого потока может быть использован свой кодек.

3.2. Работа Windows Media Player

При открытии AVI-файла, Windows Media Player сначала обращается к заголовкам потоков и просматривает их FOURCC. После чего обращается к системе кодеров, и если там он не находит нужный кодек, то обращается к веб-ресурсу (microsoft.com) откуда и загружает нужный кодек.

3.3. Кодеки

Мы несколько раз обращались к этому понятию, но до сих пор не дали ему определения. Исправим это упущение. **Кодек** — некая программная компонента, которая вызывается плеерами или программами видеомонтажа и представляет из себя или кодировщик, или декодировщик, или и то и другое. Само слово кодек является комбинацией слов «кодировщик-декодировщик»⁴.

В формате AVI однозначно описываются с помощью поля заголовка потока FOURCC.

3.4. Чего нет в AVI

- прозрачности;
- индексации;
- предпросмотра;
- временных меток;
- функции окна в окне (работа с масштабами изображения).

²Resource Interchange File Format.

³Four Character Code.

⁴Codec — coder-decoder, соответственно в английском языке.

4. Формат QuickTime

Все чего нет в AVI, можно найти в формате QuickTime.

QuickTime — проприетарная технология компании Apple, разработанная в 1991 году для воспроизведения цифрового видео, звука, текста, анимации, музыки и панорамных изображений в различных форматах.

4.1. Что есть

- группировка потоков;
- прозрачность (α -канал);
- Механизм ссылок (видео может быть распределено по нескольким файлам);
- индексация;
- предпросмотр;

- временные метки;
- функции окна в окне (работа с масштабами изображения);
- возможность работы с высотой, шириной, областью отсечения.

Последняя возможность реализована с помощью трехмерной квадратной матрицы:

$$\begin{pmatrix} a & b & 0 \\ c & d & 0 \\ dx & dy & 1 \end{pmatrix}$$

Коэффициенты a, b, c, d — элементы матрицы преобразования. Коэффициенты dx, dy — сдвиги.

Заметим, что при прочих равных условиях формат QuickTime является удобнее для использования.

5. Формат телевидения

Как было сказано выше, для телевидения размер кадра составляет 720×576 пикселей. Кроме того цвет кодируется в системе YUV.

- Y — яркость;
- U и V — цвето-разностные компоненты, это хроматический синий и хроматический красный, их иногда обозначают как C_B и C_R .

Первые телевизоры появились в Лондоне в 32 году. С появлением цветных телевизоров использование системы YUV было обусловлено обратной совместимостью с черно-белыми, в которых присутствовала только яркостная компонента.

Для преобразования RGB в YUV используют следующие формулы:

$$\begin{aligned} Y &= 0,299 \cdot R + 0,587 \cdot G + 0,114 \cdot B & [0; 255] \\ U &= 0,564 \cdot (Y - B) & [-128; +128] \\ V &= 0,713 \cdot (Y - R) & [-128; +128] \end{aligned}$$

Формулы для обратного преобразования:

$$\begin{aligned} R &= Y + 1,14 \cdot V \\ G &= Y - 0,395 \cdot U - 0,581 \cdot V \\ B &= Y + 2,032 \cdot U \end{aligned}$$

6. «Говорящая голова»

Для изучения и описания алгоритмов сжатия видео используется так называемая «говорящая голова»⁵. Это изображение некоторого персонажа, схематично изображающего человека. В зависимости от рассматриваемой ситуации «говорящая голова» совершает некоторые действия. В результате этих действий меняется картинка на экране. А по картинке определяются приемлемые (в этой ситуации) методы сжатия видео.

Вообще, на роль «говорящей головы» подойдет любой телеведущий, но для упрощения чаще используют схематично нарисованного человека.

Например, если персонаж не двигается а только говорит, то нет смысла кодировать каждый кадр целиком. Можно распознать активно меняющуюся область (лицо, губы) и от кадра к кадру изменять только ее.

6.1. Межкадровое кодирование

Рассмотренный выше пример с кодированием только отдельной части каждого кадра является примером межкадрового кодирования. Подобный метод используется в форматах FLV и FLC. Старый формат CineRac тоже использовал такой метод кодирования, но при этом еще происходило огрубление

кадров, и от кадра к кадру накапливалась ошибка. Для компенсации этой ошибки в ряд сжатых кадров добавлялся контрольный (или ключевой) несжатый кадр.

А что если персонаж, кроме того что говорит, еще и поднял руки? Тут возникают различные варианты. Первая мысль — как-то скомпенсировать движение — как это было в случае разговора персонажа. Но в реальных приложениях отслеживать перемещение объектов по экрану достаточно сложно, потому разбивают видимую область на блоки, и идея компенсации применяются уже для блоков. Реализуется она с помощью векторов смещения.

6.2. Векторное кодирование

Если за персонажем находится однородный фон, то тоже нет смысла его полностью кодировать. Эта идея называется *векторной квантизацией*. Развивая ее можно придти к мысли, что «говорящую голову» удобнее представлять векторно, в виде сплайнов.

Существует большое количество различных подходов и методов кодирования видео. Некоторые из них будут рассмотрены ниже.

7. Требования приложений к алгоритмам сжатия

- 1) Высокая степень компрессии.
- 2) Высокое качество изображения.
- 3) Высокая скорость компрессии.
- 4) Высокая скорость декомпрессии.
- 5) Устойчивость к ошибкам передачи.
- 6) Учет специфики исходного файла.
- 7) Низкая стоимость аппаратной реализации.
- 8) Редактируемость.

8. Критерии сравнения

- 1) Худший, средний, лучший коэффициенты сжатия.
- 2) Классы изображения или видео на которые ориентирован алгоритм.
- 3) Оценка потерь качества.
- 4) Симметричность — отношение выбранной характеристики кодера, к аналогичной характеристике декодера.
 - по времени
 - по объему памяти.

⁵В оригинале — talking head.

9. Монтажные эффекты и видео переходы

9.1. Таимкоды

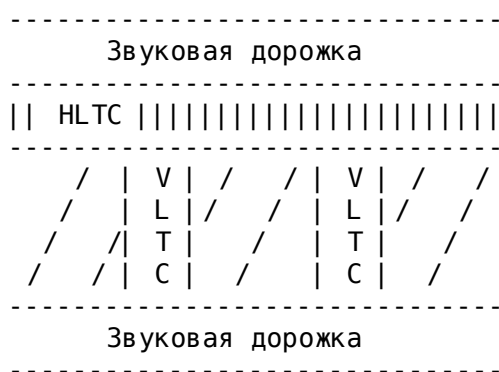
Это логическая запись — позволяет идентифицировать точку в видео ряде. Вид таимкода: ЧЧ.ММ.СС.КК (посчитать сколько бит) Максимальный вариант: 99.59.59.XX, XX — зависит от используемого формата (PAL KK = 25).

9.2. Ленты

Нас пока не интересует цифровая у нас запись или аналоговая

Все становится очень сложно в случае с лентной протяжкой. В этом случае используется :

- HLTC — horizontal length time code.
- VLTC — vertical length time code.



VLTC пишется в межкадровое пространство.

9.3. Способы монтажа видео переходов

Здесь речь идет исключительно о переходах.

- Cut — прямая склейка, переход без нахлеста;
- Fade — затухание (до черного экрана);
- Dissolve — растворение;
- Wipes — шторы, герелистывание страниц.

Все остальные эффекты являются сочетанием этих четырех базовых блока. После создания нарезки ее нужно экспортировать в EDL (Edit Decision List). Это между народный формат. Его понимает большинство продуктов.

9.4. Видео монтажные приемы

Chroma/Luma Key — ключ для объявление части поверхности ролика прозрачной. Самый распространенный пример. Используем «говорящая голова». Сажаем диктора на монотонный фон (обычно синий или зеленый, берут цвета максимально контрастные с цветом тела человека). «Сключивавшись» по этому фону, его объявляют прозрачным и подставляют на его место некоторое изображение. Экран должен быть безшовный — шва не должно быть.

Так же используется Риф-проекция.

9.5. Виртуальные студии

Обсчет в масштабе реального времени погружением реальных объектов в виртуальную среду. Например есть человек, которого снимают на синем или зеленом фоне. Камера имеет ряд датчиков. Датчик на трансфокаторе (Трансфокатор — устройство меняющее фокус.) и на штативе. (Это очень дорогие датчики) Сигнал передается кластеру из очень производительных компьютеров. Компьютер обсчитывает задник (background) и передник (foreground). Причем если виртуальный объект помещен перед реальным актером, то необходимо учесть тень этого объекта на актера. Здесь используется Chroma Key.

9.6. Виртуальные актеры

Виртуальный актер — человек всегда остающийся за кадром. Это люди которые сообщают искусственному персонажу мимику, движение звук, передает характерные особенности движения волосного покрова. Самый первый полнометражный фильм, где за каждым искусственным актером существовал, виртуальный актер — Finall Fantasy. Как снимают артикуляцию? — Это очень сложно. На лицо человеку наклеивают свето-возвращающий элемент, потом накладывают данные трехмерную сетку. Для движения вариантов больше:

- свето-возвращающие датчики;
- проводные датчики;
- беспроводные датчики.

Пропуск

10. Алфавитные методы

Подходят к процессу сжатия с точки зрения алфавита. Например:

- Азбуки Морзе;
- Азбуки Брайля (перевод зрительных образов в тактильные)

Мы рассмотрим метод Хаффмана (Huffman). Пусть задан алфавит A состоящий из конечного числа букв.

$$A = \{a_1, \dots, a_R\}$$

Конечную последовательность символов, из этого алфавита будем называть словом, а число n длиной слова.

$$W_A = a_{i_1}, a_{i_2} \dots a_{i_n}$$

$$l(W_A) = n$$

Пусть \exists алфавит B

$$= \{b_1, \dots, b_Q\}$$

В нем существуют слова:

$$W_B = b_{i_1}, b_{i_2} \dots b_{i_n}$$

И пусть $S(B)$ — множество не пустых слов алфавита B

Пусть задано отображение F которое каждому слову из алфавита A ставит соответствие слово из алфавита B .

$$F : W_A, W_B \in S(A) \mapsto W_B$$

$$W_B = F(W_A), W_B \in S(B);$$

W_B — код слова W_A Например: В азбуке Морзе

- W_A — русский алфавит $l(A) = 33$;
- W_B — точка и тире $l(A) = 2$;
- F — таблица.

Причем F имеет вид:

$$\begin{array}{ll} a_1 & \rightarrow W_{B_1} \\ a_2 & \rightarrow W_{B_2} \\ & \dots \\ a_R & \rightarrow W_{B_R} \end{array}$$

Если:

$$W_A = a_{i_1}, a_{i_2} \dots a_{i_n} \Rightarrow W_B = W_{B_{i_1}} W_{B_{i_2}} \dots W_{B_{i_n}}$$

Если n совпадает, то кодирование называется **алфавитным**. В нашем случае алфавит состоит из 0 и 1.

$$B = \{0, 1\}$$

Рассмотрим пример «ЁЁЁЁУУУ». Вероятности

$$\begin{aligned} \text{Ё} &= \frac{4}{6} = 0 \\ \text{У} &= \frac{2}{6} = 1 \end{aligned}$$

«ЁЁЁЁУУУО» Вероятности:

$$\begin{aligned} \text{Ё} &= \frac{4}{7} = 0 \\ \text{У} &= \frac{2}{7} = 10 \\ \text{О} &= \frac{1}{7} = 11 \end{aligned}$$

Важно обязательно уметь однозначно восстанавливать.

Пусть наше слово состоит из двух частей. Тогда первую часть называют префиксом слова.

Пусть $W_b = W'_B W''_B$

- W'_B — префикс слова
- W''_B — конец слова

Схемы F обладает свойством префикса если

$$\forall i, j (1 \leq i, k \leq R)$$

слово W_B не является префиксом W_B

Теорема: Если схема F обладает свойством префикса, то алфавитное кодирование будет взаимно однозначным.

$$\exists P_1, \dots, P_R (\sum_i P_i = 1);$$

Вероятность появления a_i в сообщении W_A . Для каждой схемы F можно посчитать некое среднее

$$e_{\text{cp}} = \sum_i^R P_i l_i, l_i = l(W_i)$$

$$\exists e^* = \min_F l_{\text{cp}}$$

F^* — код Хаффмана

10.1. Алгоритм построения кодов Хаффмана

Считаем $B = \{0, 1\}$:

- 1)
 - Упорядочиваем все буквы входного алфавита A в порядке убывания появления в сообщении.
 - Считаем все соответствующие слова W_B из алфавита B пустыми.
- 2)
 - Обедняем 2 последних символа $a_{i_{R-1}}, a_{i_R}$ с наименьшими вероятностями. $P_{i_{R-1}}, P_{i_R}$ в 1 псевдосимвол. $a' a_{i_{R-1}}$, с вероятностью a_{i_R} с $P_{i_{R-1}} + P_{i_R}$
 - Дописываем 0 в начало слова $W_{B_{i_{R-1}}}$ ($W_{B_{i_{R-1}}} = 0W_{B_{i_{R-1}}}$) и 1 в начало $W_{B_{i_R}}$
- 3)
 - Удаляем из списка упорядоченных символов наши 2 символа.
 - Заносим a' с суммарной вероятностью. Список остается отсортирован по вероятности.
 - Проводим заново шаг 2. Добавляя при необходимости 1 или 0. Для всех слов W_i соответствующих всех символов до тех пор пока не останется 1 псевдосимвол.

Пример: Пусть алфавит состоит из

$$A = \{a_1, a_4\} R = 4$$

$$P_1 = 0.5$$

$$P_2 = 0.24$$

$$P_3 = 0.15$$

$$P_4 = 0.11$$

И мы помним, что $\sum p = 1$. Далее:

1)

$$\begin{array}{ll} P_1 = 0.5 & P_1 = 0.5 \\ P_2 = 0.24 & P_2 = 0.24 \\ P_3 = 0.15 & P_3 = 0.15 \rightarrow 0 \\ P_4 = 0.11 & P_4 = 0.11 \rightarrow 1 \end{array}$$

2)

$$\begin{array}{ll} P_1 = 0.5 & P_1 = 0.5 \\ P_2 = 0.24 & P_2 = 0.24 \rightarrow 0 \\ P_3 = 0.26 & P_3 = 0.26 \rightarrow 1 \end{array}$$

3)

$$\begin{array}{ll} P_1 = 0.5 & P_1 = 0.5 \rightarrow 1 \\ P_2 = 0.5 & P_2 = 0.5 \rightarrow 1 \end{array}$$

Таким образом

$$\begin{aligned}a_1 &= 0 \\a_2 &= 11 \\a_3 &= 100 \\a_4 &= 101\end{aligned}$$

$$l(W_A) = 100 \text{ Байт} \quad l(W_B) = 176 \text{ Бит}$$

Минусы алгоритма Хаффмана:

- надо сохранять схему F что увеличивает битовый бюджет;
- двухпроходность;
- на 1 проходе собираются вероятности.

Сама схема адаптированная. Существуют коды хаффмана с фиксированной таблицей кодов. Например для английского текстов существуют коды фиксированные схемы. $e == 0.13$, $z == 0.024$

10.2. Коды Хаффмана с фиксированной таблицей

Факсимильное сжатие: В ITU-T взяли 8 эталонных документов:

- 1) Типичное деловое письмо на английском;
- 2) Рисунок электрической цепи от руки;
- 3) Печатная форма заполненная на машинке на французском.
- 4) Плотный напечатанный отчет на французском.
- 5) Техническая статья с иллюстрациями на французском.
- 6) График с напечатанными подписями на французском.
- 7) Произвольный плотный документ.
- 8) Рукописная записка белым по черному на английском.

Они предложили стандарты передачи факса $T1, T2, T3, T4$. Факс машина сканирует документ по строкам. Переводя каждую строку в последовательность черных и белых точек называемых пелами.

- Горизонтальное разрешение 8.05 пелов / мм
1728 пелов на 8.5";
- Вертикальное разрешение 3.85 либо 7.7 линий / мм
(978 / 1956).

Чаще всего встречаются:

- 2, 3, 4, черных пелов.
- 2 - 7 белых пелов.

Длинна серии белые коды черные коды

```
-----
0 00110101 0000110111
1 000111 010
2 0111 11
3 1000 011
4 1011 0011
5 1100 0010
6 1110 0010
7 1111 00011
8 10011 000101
... ..
63
```

T3 использует модифицированные коды хаффмана. добавляют RLE.

Длинна серии белые коды черные коды

```
-----
64 11011 0000001111
128 10010 4x0112x013x0
192 ... ..
```

Теорема Шенона: Энтропия символа A имеющего вероятностью F — количество информации в A. $= -P \log_2 P$

Построить график.!!!

Коды хаффмана достигают энтр пределов, если вер символов кратно 2

11. Алфавитные методы

1) [0,1);

2) 2.1 разделить текущий интервал пропорционально вероятности каждого символа. 2.2 выделить подынтервал соответствующий текущему символу и назначить его новым текущим интервалом. 3) Когда весь текущий файл будет обработан, выходом алгоритмом объявляется любая точка внутри текущего интервала.

$$a_1 a_2 a_3$$

$$P_1 = 0.4 P_2 = 0.5 P_3 = 0.1$$

$$a_2, a_2, a_2, a_3;$$

$$[0; 0.4) [0.4; 0.9) \Rightarrow [0.6; 0.85) \Rightarrow [0.7; 0.825) \Rightarrow [0.8125; 0.8250) [0.9; 1)$$

Декодер так же как кодер режет весь интервал на 3 отрезка

Этот алгоритм симметричен. Симметричность == 1;

11.1. Коды грея посмотреть

Посмотреть самим

12. Алгоритмы вероятностного кодирования

Посмотреть самим

13. Физиологические предпосылки для методов сжатия с частичной потерей качества оригинала

- Яркостная информация воспринимается человеком с больше детализацией чем световая.
- Так как глазная сетчатка имеет существенно больше рецепторов, чем отходящих от них нейронов, то информация от групп нейронов передается сразу от групп рецепторов. Это особенность мешает человеку скопление мелких предметов от единого целого.
- Обработка яркости и контуров объектов производится разными полушариями мозга и с разными скоростями. Контуров объектов можно представлять с меньшей точностью.
- Инерционность зрения. Это инерционность фотохимических процессов. инерционность распространения сигнала по нейронным сетям. Эффект *послевидения* — принцип построения временной избыточности.

14. Законы субъективного восприятия яркости

Закон Тальбота. Если L — яркость экрана

$$L_{cp} = 1/T \int_0^T L(t) dt$$

L — яркость экрана.

$T = 1/f$; f — частота мельканий.

Этот закон позволяет нам не видеть момент смены кадров на ленте.

14.1. Закон Блоха-Шарпантье

Ориентирован на импульсную перерисовку экрана.

$$LT = const$$

$T_u = 60 - 100$ — период когда экран светится

15. Метод оценки качества

Методы оценивают, как статичные изображения так и оценки видео. Большая часть методов работает с яркостями (90 %).

Все методы делятся на 2 большие группы:

- объективные; это повторяемые методы.
- субъективные, экспертные; предсказуемость этих методов страдает.

Часто вынуждены работать с субъективными методами, ибо все работает для людей. и человек мера всех вещей.

15.1. Субъективные методы оценки качества

Существует несколько стандартов. Субъективные методы:

- Шкала деградации с двумя стимулами — Double Stimulus Impairment Scale — DSIS.
- Непрерывная оценка качества с одним стимулом.
- Попарное сравнение.

Требования к экспертам:

- размер группы от 4 до 40;
- *одинаковая острота зрения и нормальное цветовосприятие;*
- эксперты не должны быть профессионалами в области оценки качества.

15.1.1. Сложности тестирования

- дистанция наблюдения; индивидуальные предпочтения расстояния до экрана у всех разное. существуют общие рекомендации: *для объектов чтения:* должны быть на расстоянии вытянутой руки; *для телевизоров:* 3-4 диагонали.
- индивидуальные интересы и ожидания;
- тип дисплея и его характеристики;
- условия наблюдения: освещенность, комфортность по температуре.
- точность воспроизведения; точность цветопередачи и контрастность.
- сопутствующий саундтрек.

Стандарт ITU-R BT500, «Методика субъективной оценки качества телевизионного изображения». Впервые опубликована в 1974 году. — Конкуренты VQEG.

15.1.2. Шкала деградации с двумя стимулами

Ролик должен простимулировать эксперта. Участники просматривают. Эталонную сцену а затем ту же сцену после ее обработки в тестируемой системе.

Каждый участник оценивает деградацию изображения сцены после обработки. одним числом по пятибальной шкале:

- 1 — очень раздражает
- 2 — раздражает
- 3 — слегка раздражает
- 4 — заметно не раздражает
- 5 — не раздражает

Баллы выставляются для некоторого количества различных сцен. Тут важно досматривать сцены до конца. И при показе нескольких вариантов одного и того же ролика — важно перед каждым показом сжатого файла показать оригинал.

15.1.3. Непрерывная оценка качества с одним стимулом.

Зритель непрерывно оценивает одну телепрограмму длительностью 10-20 минут, а не серию тестовых сцен. Зрительные рейтинги записываются несколько раз в минуту. В этом методе зритель видит только сжатый ролик.

15.1.4. Попарное сравнение

На экране одновременно, 2 стимула, а пары создаются путем всевозможной комбинации из эталона и тестируемых методов. При этом, эксперту не сообщается какой из роликов эталон.

A — исходный *B*, *C* — тестируемый ролики

Пары: *AB*, *BC*, *CB*, *BA*, *AC*, *CA*.

15.2. Дополнительные размерности оценок

Никто не запрещает эксперту дать дополнительную оценку:

- 1) яркости;
- 2) контраста;
- 3) цветопередачи;
- 4) четкость контуров и ореол вокруг них;
- 5) стабильность фона;

- 6) плавность перемещения объектов сцены;
- 7) дрожание;
- 8) эффект муара — размывание;
- 9) эффект снега или комаров;
- 10) двоение изображения и тени.

Видео по телику передается в MPEG-2.

15.3. Объективные методы

Объективные методы:

- $\sup(A - B)$ Функция показывает значение максимальной разности.
- разность мат. ожиданий Покажет на сколько в среднем отличается яркость картинок. На выходе число, которое характеризует изменение яркостью
 - разность мат. ожидания
 - мат. ожидание разностей

- СКО =

$$\frac{1}{n}(M_1 - M_2)^2$$

n — число пикселей

- PSNR — Peak Signal to Noice Ratio

$$PSNR = 20 \log_{10} \frac{MAX * M * N}{\sqrt{\sum_{m,n} (A_{i,j} - B_{i,j})^2}}$$

MAX — пик сигнала M, N — размерности. $A_{i,j} - B_{i,j}$ — попиксельная разница.

Подумать как это связано с СКО.

15.4. Объективные методы

Подумать как это связано с СКО. \log — ?

$$\sqrt{\sum_{m,n} (A_{i,j} - B_{i,j})^2} \text{ --- ?}$$

MAX = 255

PSNR измеряют в дБ.

15.4.1. Психофизиологическая кривая

Связь психологического качества от PSNR.

16. Контекстные методы

16.1. JND

JND — метод едва заметны различий. Just Niticeable Diff

Разработана компанией Sarnoff (очень крупная Американская компания, которая занимается разработкой видеоборудования) в середине девяностых. Сарнов сам по себе советский ученый, который в 30-х годах уехал в США. Он приложил руку к созданию первого телевизора в СССР, и первого цветного телевизора в США.

Опишем как она будет работать.

- Фильтрация изображений
 - Разделение на массивы яркости и цветности
 - Определение средней яркости
 - Три отдельных измерения контраста:
 - В терминах ориентации, мерцания и цветового смещения.
- Маскирование энергии контраста Энергия контраста — абсолютные значения разницы контраста не влияет
- Учитывается понижение чувствительности зрительной системы человек к искажениям в сложных областях изображения.
- Построение карт JND. На этом шаге строится 2 карты, Одна для заметных различий по яркости и для заметных различий по цвету. Чем заметнее изменение тем ярче соответствующая область на карте. Это весовые матрицы для определения артефактов.

Преимущества: Более устойчивая зависимость с субъективными методами чем психофизическая кривая.

Недостатки: Патентована была.

JND была даже реализована на аппаратном уровне в Tektronix PQA-200 — аппаратный измеритель качества сигнала.

16.2. PDM

PDM — метрика воспринимаемых искажений — Perceptual Distortion Metric

Строится карта. Автор — Стефан Винкбез разработана в начале двухтысячных Используется метод изотропического контраста.

Мы тут строим несколько уровней контраста.

16.3. SSIM

SSIM — метрика структурного сходства — Structural SIMilarity

Автор — Джун Венг, и т.д.

Строится карта. Работает с локальным анализом изображений в терминах структурного сходства. Должен быть пространственный анализ структура изображения. На выявление где у нас будут виднее артефакты. На эту тему можно почитать www.compression.ru

17. По курсовой работе

17.1. Реферат

6-8 страниц используемые методы сжатия.

При защите можно побеседовать по теории.

17.2. Вычислительная часть

Взять по три кадра исходных (без потерь): титры быстрое движение отснятое видео

И по три кадра для каждого из трех кодеков. С помощью утилиты можно работать в сумме по всем кадрам.

Для курсовой нужно обязательно использовать. линейная разность, м.о, PSNR (картинка)

Что делать: Инвертировать и линейную разность высчитывать по контрастности можно утилиты писать самим или использовать утилиту с сайта www.compression.ru

17.3. Аналитическая часть

Продемонстрировать свои аналитические навыки

Что нужно:

- Сравнение кодеков между собой

- Операться на эксперивенальнцю часть
- Оценка должна быть комплексной (не только качества). Опираься на факты.
- сравнение с субъективными оценками (из лр 4)
- главное хорошее обоснование.

Методы сжатия с потерями

18. VQ - Метод векторной квантизации

Устраняет пространственную избыточность Сущ. 2 вида:

spatial temporal

Идея — разбиение кадров на блоки и замена похожих блоков на 1.

Вводится некоторый критерий качества Q . Пространство разбивается на участки.

$$d(i, j) < Q.$$

то i и j указывается блок.

По тем блокам которые останутся разбиваем еще раз на блоки.

Практика показывает что для размеров блоков 2 на 2 пиксела — всегда работает. Такая поисковая таблица — книга шифров.

19. Временное квантование

I 255 253 210 249 II 254 253 230 250 D -1 0 20 1 скалярное квантование 0 0 20 0 1 + 2 ===
CINERAK

(15-й кадр ключевой.)

20. Контурное кодирование

Представляет кадры как контуры ограничивающие текстурированную область. Это метод пространственное сжатие.

Алгоритмы выделения контуров. Контуры кодируются как контрольные точки сплайнов. А текстуры компрессируются отдельно.

Этот стандарт хорошо подходит для видео конференций: H.261 H.263 H.264

21. Цветовое прореживание

Скажем про стандарты: PC 601:

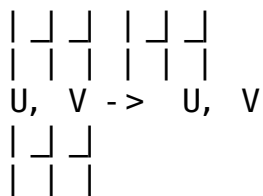
$$RGB \rightarrow YUV/PC601/$$

Диапазон $[0, 255]$; $[0, 255]$, $[-128, 128]$

Кроме, есть еще, и REC 601 (TV): Диапазон $[0, 255]$; $[16, 255]$, $[16, 240]$, $[16, 240]$,

Человеческий глаз наиболее чувствителен для изменения яркости больше.

$Y \rightarrow Y$



а) 4:4:4



б) 4:2:2



в) 4:2:0



г) 4:2:0



Цветовое прореживание используется JPG (Joint Print Exp Group).

Видео составленное из картинок сжатых по дипегу М-JPEG (moution-JPEG).

Применяется с высокими яростной и цветовой разрешениями. На схемах его использо-
вать не эффективно и даже опасно. I) Производится преобразование RGB-YUV II) Цве-
товое прореживание — 4.2.2 или 4.1.1 III) Яростная плоскость разбивается на блоки (8

на 8) (16 на 16) Прореженные цветоразностные плоскости. IV) каждый блок подвергается дискретному кос-преобр.

$$C_{i,j} = \sum_{n=0}^{N-1} \sum_{m=1}^{M-1} C(i)C(j) \cos\left(\frac{(2n+1)i\pi}{2N}\right) \cos\left(\frac{(2n+1)j\pi}{2M}\right) f_{n,m}$$

$$(i) = 1/\sqrt{(2)}, i = 0; 1, i \neq 0$$

V) полученну матрицу кожф. поэлементно делим на матрицу квантований.

$$A_{n \times n} = a_{i,j} = 1 + \alpha(i + y), N - 1, i, j = 0$$

α — коэф. качетсва. VI) Зигзаг сканирования. Разворачиваем матрицу в строку. VII) Полученная последовательность сжимается по RLE четвертого типа. VIII) Либо Хаффман либо арифметическое кодирование.

Один из паразитарных эффектов — эффект мазайки. Эффект Гиббса — визуально в за-
двоенных контурах.

21.1. JPEG

Ключевой стандарт. Прошлый раз прошлись по основным этапам. Рассмотрели те шаги на которых возникают основные шаги потери качества. А именно эффект Гиббса.

Группа MPEG-методов

Мы будем говорить про видео-последовательности.

MPEG — motion picture expert group.

- MPEG-1 — видео и аудио;
- MPEG-2 — видео и аудио;
- MPEG-3 — аудио;
- MPEG-4 — видео и аудио;

22. MPEG-1

Ориентирован на системы записи компакт-диски CD-ROM и низкоскоростные каналы передачи телевизионных изображений со скоростью до полутора мегабит в секунду или меньше.

Используется стандарт разретрки в четкостью в 4 раза меньше чем у вещательного теле: 288 активных строк и 352 отсчета в активной части строки. Этот формат используется для Video-CD со скоростью 1х.

На дом посчитать пропускную способность однокоростного потока для Audio-CD.

23. MPEG-2

Этот стандарт используется для передачи цифрового телевещания.

- DVD (*.vob);
- DVB (digital video bloodcosting — спутниковая картинка).

Промышленное вещание было начато в 1995 году. Нету одного стандарта MPEG-2.

MPEG-2 — это семейство согласованных, совместимых цифровых std.

Посмотреть книжку Артюшенко «Цифровое сжатие видеоинформации и звука».

В MPEG-2 существуют видеопрофили. С одной стороны они привязаны к пространственному разрешению.

Одно измерение:

- CIF (Common Intermediate Format) — 352 x 228;
- QCIF – 176 x 144;
- 4CIF — 704 x 576 (720 x 576);
- 16CIF — 1408 x 1152 (1440 x 1152);

Другое измерение:

- простой профиль без Б-кадров 4:2:2 — 4 Мб/с;
- основной профиль без Б-кадров 4:2:0 — 15 Мб/с;
- профиль с масштабируемым отношением Сигнал-Шум и Б-кадрами 4:2:0 — не предусмотрено;
- высший профиль с Б-кадрами 4:2:0 — 80 Мб/с;
- высший профиль с Б-кадрами 4:2:2 — 80 Мб/с;

4:2:0: Y: 0 1 2 3 U: 4 — по первому столбцу первой строки V: 5 — по первому столбцу первой строки 4:2:2: Y: 0 1 2 3 U: 4 — по первому столбцу 6 — по первому столбцу V: 5 — по первому столбцу 7 — по первому столбцу 4:4:4: Y: 0 1 2 3 U: 4 8 6 10 V: 5 9 7 11
Есть еще 4:1:1 — это DV-video. Но в Mpeg в явном виде не используется.

24. Временная избыточность

GOP (Group of Pictures) как правило состоит из 12 кадров.

порядок воспроизведения [I][B][B][P][B][P][B][P][B][P][B][B]

I — интро, опорный кадр, это базовое и самодостаточное при декодировании изображение по JPG. Степень сжатия минимальна.

P — (predicted) предсказанный, восстанавливается с участием предыдущего P или I кадра, кодирование осуществляется с компенсацией движения и предсказания вперед. Тут (MPEG-2) только блочное. P-кадр ссылается на предыдущий I кадр.

B — (bidirectionally predicted) двустороннее предсказывание. B-кадр восстанавливается с участием предыдущего и/или последующего I или P кадра. Степень сжатия максимальна. 4 варианта:

- 1) Предсказание вперед по последнему ближайшему I или P кадру.
- 2) Предсказание назад по последующему ближайшему I или P кадру.
- 3) Двухнаправленное предсказание
- 4) Внутрикадровое сжатие без компенсации движения (по JPG)

порядок передачи и хранения: [I0][P3][B1][B2][P6][B4][B5]

MPEG-2 (I) = M-JPEG используется в видеокамерах.

Для MPEG-2 должен существовать буфер (линия) задержки.

25. Компенсация движения

Методы поиска смещенных объектов. В MPEG-2 используется метод основанный на макроблоках.

Разбиваем всю картинку на блоки.

25.1. Метод прямого сравнении блоков изображении

Вычислении нормы разности между заданным блоком и его смещением в другом кадре на каждый вектор из области поиска векторов смещения. Достоинство: Точность. Недостаток: Скорость.

26. Виды монтажа

26.1. линейный

26.2. нелинейный

Произвольный доступ.

26.3. гибридный

27. Методы субоптимального поиска

Методов много.

- 1) Сигнатурные методы.
- 2) Поиск с разбавленным расстоянием.
- 3) Локализованный телескопический поиск.
- 4) Двумерный логарифмический поиск.
- 5) Многомерный поиск.
- 6) Метод корреляции фаз.

27.1. Сигнатурные методы

Несколько шагов, сокращая количество кандидатов на каждом шаге. На первом шаге. Все кандидаты проверяются с помощью простой, но быстрой метрики — меры расходимости. На следующих шагах оставшиеся кандидаты проверяются с помощью более сложных и точных метрик.

27.2. Поиск с разбавленным расстоянием

x — блок. Медленные объекты — более четкие. Быстрые объекты — смазанные.

```
-----  
x  x  x  x  x  x  x  x  
x  x  x  x  x  x  x  x  
x  x  x  x  x  x  x  x  
x  x  x  x  x  x  x  x  
x  x  x  x  x  x  x  x  
x  x  x  x  x  x  x  x  
x  x  x  x  x  x  x  x  
-----
```

Берем соседей с шагом 1, 3, и т.д И смотрим метрику от кадра к кадру.

```

-----
x (x) x x (x) x x (x)
x x x x x x x x
x x x (x)(x)(x) x x
x (x) x (x)(X)(x) x (x)
x x x (x)(x)(x) x x
x x x x x x x x
x (x) x x (x) x x (x)
x x x x x x x x
-----

```

(X) — центральный.

27.3. Локализованный телескопический поиск

Группа методов, опирается на известные значения векторов смещения полученные для предыдущего кадра.

```

-----
x (x) x (x) x (x) x (x)
x x x x x x x x
x (x) x (x)(x)(x) x (x)
x x x (x)(X)(x) x x
x (x) x (x)(x)(x) x (x)
x x {x} x x x x x
x (x) x (x) x (x) x (x)
{x} x x x x x x x
-----

```

Объект больше блока. Несколько блоков двигаются в 1 направлении. Вектор смещения того же блока в предыдущем кадре.

27.4. Двумерный логарифмический поиск

- 1) Параметр D , на основании которого задается размер шага. $[\cdot]$ — взятие целой части.

$$S = 2^{\lceil \log_2 d \rceil - 1}$$

$$(a, b), (a, b + s), (a, b - s), (a + s, b), (a - s, b)$$

- 2) Выбирается лучший. Пусть это (x, y) , если $(x, y) = (a, b)$ то $s = \frac{s}{2}$, иначе, $s = s$, $(a, b) = (x, y)$, идем в 1.
- 3) Если $s = 1$, то выбираем самый лучший из 9 соседей.

Пример: Первый шаг

$d = 8;$
 $s = 4;$

```

-----
x  x  x  x  (x) x  x  x  x
x  x  x  x  x  x  x  x  x
x  x  x  x  x  x  x  x  x
x  x  x  x  x  x  x  x  x
(x) x  x  x [ (X) ] x  x  x  (x)
x  x  x  x  x  x  x  x  x
x  x  x  x  x  x  x  x  x
x  x  x  x  x  x  x  x  x
x  x  x  x  (x) x  x  x  x
-----

```

Второй шаг

```

-----
x  x  x  x (1) x  x  x  x
x  x  x  x  x  x  x  x  x
x  x  x  x {2} x  x  x  x
x  x  x  x  x  x  x  x  x
(1) x {2} x (1) x[{2}]x (1)
x  x  x  x  x  x  x  x  x
x  x  x  x {2} x  x  x  x
x  x  x  x  x  x  x  x  x
x  x  x  x (1) x  x  x  x
-----

```

Третий шаг

```

-----
x  x  x  x (1) x  x  x  x
x  x  x  x  x  x  x  x  x
x  x  x  x {2} x (3) x  x
x  x  x  x  x  x  x  x  x
(1) x {2} x (1) x[{2}]x (1)
x  x  x  x  x  x  x  x  x
x  x  x  x {2} x (3) x  x
x  x  x  x  x  x  x  x  x
x  x  x  x (1) x  x  x  x
-----

```

Конец

```
-----  
x  x  x  x (1) x  x  x  x  
x  x  x  x  x  x  x  x  x  
x  x  x  x {2} x (3) x  x  
x  x  x  x  x (4) (4) [4] x  
(1) x {2} x (1) (4) {2} (4) (1)  
x  x  x  x  x (4) (4) (4) x  
x  x  x  x {2} x (3) x  x  
x  x  x  x  x  x  x  x  x  
x  x  x  x (1) x  x  x  x  
-----
```

[4] — оптимальный

27.5. Многомерный поиск

Помимо поисков сдвигов ищутся вращения растяжения и сжатия. Еще имеет значение яркость. Методы поиска похожи на фрактальный алгоритм сжатия.

27.6. Метод корреляции фаз

Взаимная корреляция двух изображений. Это один из самых прогрессивных методов. Его идея — вычислив его функцию, эта функция будет показывать максимумы.

$$f_1(x, y) = f_2(x + x_0, y + y_0)$$

вектор смещения.

ПО теореме Фурье о смещении.

$$\begin{aligned} F_1(a, b) &= \\ &= \iint_{-\infty}^{+\infty} f_1(x, y) e^{-i(xa+yb)} dx dy = \\ &= \iint_{-\infty}^{+\infty} f_2(x + x_0, y + y_0) e^{-i((x+x_0)a+(y+y_0)b)} d(x + x_0) d(y + y_0) = \\ &= e^{-i(x_0a+Y_0b)} F_2(a, b) \end{aligned}$$

совместные корреляционные функции, которая имеет максимум в x_0 и y_0

28. Характерные дефекты MPEG

- 1) Шевелящийся фон — квантизационный шум.
- 2) Мозаичность
- 3) Призрачные следы, танцующие глаза. Это ошибки предсказания.
- 4) Грязное оконное стекло.
- 5) Цветные полосы.

MPEG-2 — используется в DVD.

29. MPEG-4

Подмножеством MPEG-4 является формат H.264 (десятая глава ISO MPEG4) — видео конференции.

Существует не только стандартный MPEG-4, но и microsoft MPEG. Dvix, Xdiv — кодаки, которые используют стандартный MPEG.

29.1. MPEG-4

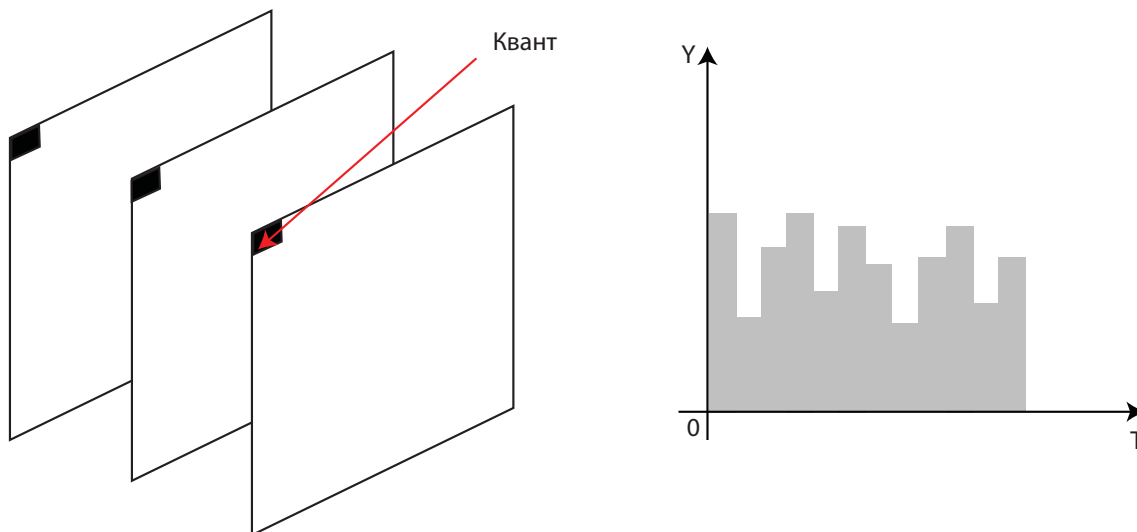
Особенности:

- 1) Кодирование объектов имеющих не правильную форму.
- 2) Можно кодировать объекты переднего и заднего плана независимо.
- 3) Кодирование объектов анимации на основе двухмерных и трехмерных полигональных сеток (анимация лиц и фигур людей).
- 4) Масштабируемое кодирование, для гибкой схемы передачи на разных битовых скоростях.
- 5) Кодирование неподвижных текстур.

Понятие VideoObject.

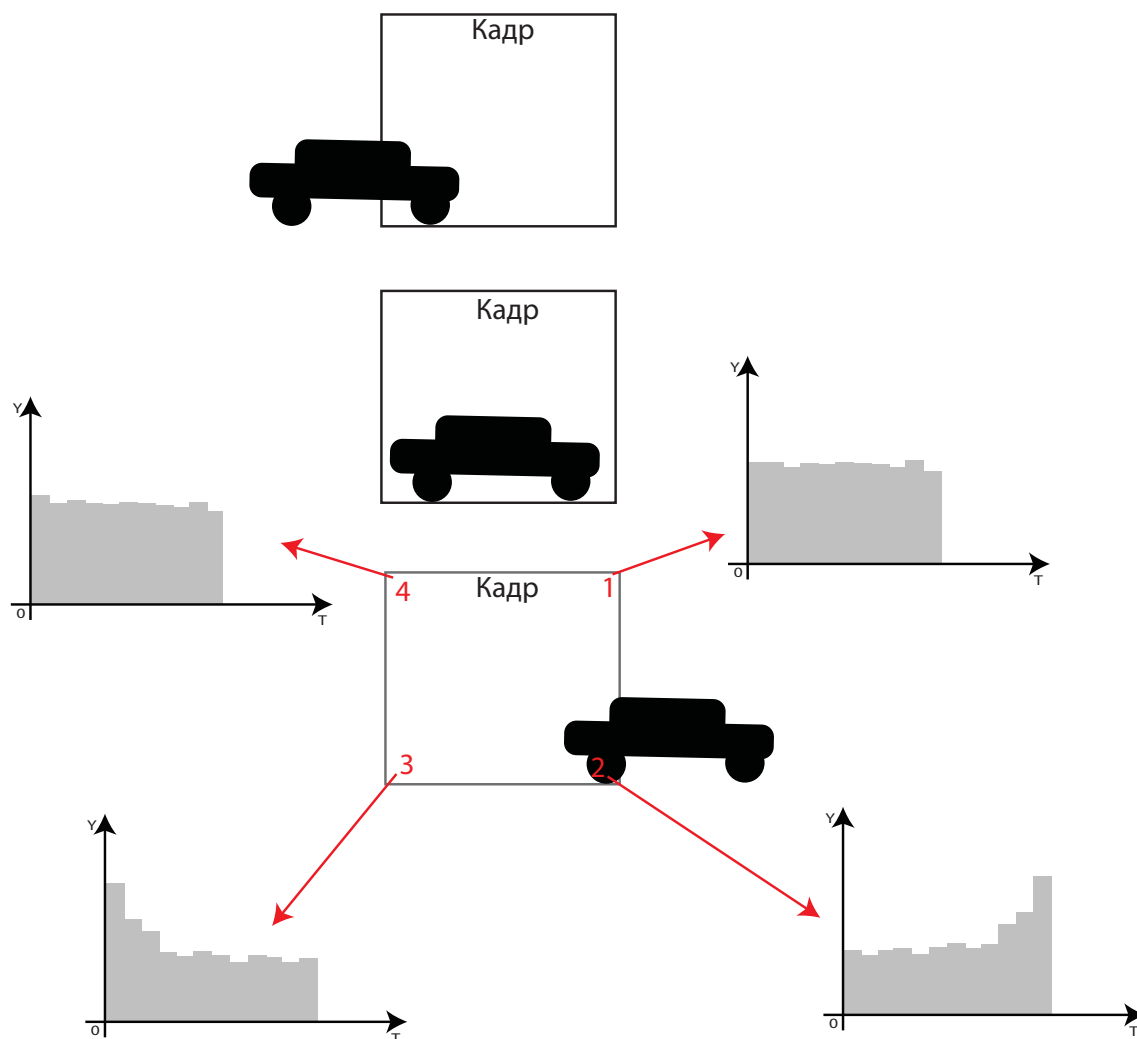
Метод кросс-кадровой интерполяции

Это не промышленный стандарт. Метод известный не очень широко. Метод работает только с видео. Чем длиннее одна сцена тем лучше работает этот метод.



Рассмотрим видео ряд в дискретном пространстве яркостей.

- 1) Исходный по ключевым кадром разбивается на сцены (ключевые сегменты).
- 2) Для каждого элементарного кванта строится гистограмма яркости от времени.
- 3) Все гистограммы специальным образом группируются в классы подобия сегмента так, что в пределах одного класса кванты ведут себя примерно одинаково.
- 4) Кванты не отнесенные ни к одному из классов подобия трактуются как одиночный незначительный шум. И заменяются при декомпрессии средневзвешенными значениями соседей по кадру.
- 5) В каждом классе ищется базисная (опорная) функция яркости. Она имеет минимальное несоответствие с остальными. Остальные функции квантов класса описываются либо сдвигами опорной функции либо в местах значительных расхождений. Запоминаются отдельно. Это «выбросы».
- 6) Выбросы между опорной функцией и остальными пространственно группируются в кадре по всем классам и сжимаются по ДКП (дискретное кос. преобразование).
- 7) посегментно заносится:
 - принадлежность квантов классам подобия;
 - базисные функции классов;
 - коэффициенты ДКП единичных выбросов.



30. Статистика распределения квантов по классам

- Фон — 50 %;
- Значимые объекты — 10-30 %;
- Детали — 20-40 %.

31. Преимущества и недостатки

Преимущества:

- возможность фильтровать случайные точечные артефакты;
- эффективное сжатие однородных данных находящихся в разных местах кадра;
- максимально эффективное интернет-вещание.

Недостатки:

- сложность и асимметричность;
- артефакты неправильной формы.

Фрактальное сжатие

Литература:

- Ватолин
- Крановер "Фракталы и хаос"

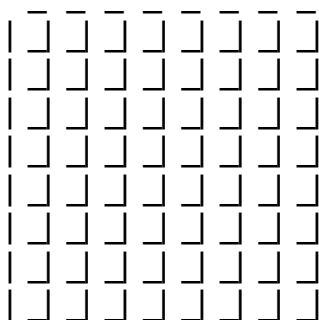
32. Построение алгоритма

Идея: это поиск самоподобных областей в изображении и определение для них параметров аффинных преобразований

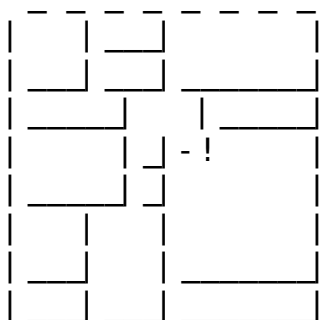
В общем случае потребуются перебор и сравнение всех возможных областей изображения разного размера.

Допущения учебного варианта алгоритма.

- все области суть квадраты со сторонами, параллельными сторонам изображения;
- все изображение равномерной сеткой разбивается на набор доменных блоков, все доменные блоки — квадраты, имеют фиксированный размер;



- все изображение покрывается ранговыми областями, которые берутся через точку по осям x и y .



возможен внахлест

- при переводе области из ранговой в доменную уменьшение размеров происходит ровно в 2 раза;
- при переводе ранговой области в доменную перевод возможен только 0, 90, 180, 270, так же допускается зеркально отражение;
- учитываем масштабирование по яркости.

Алгоритм сводится к перебору всех доменных блоков и подбору для каждого из них соответствующего ему ранговой области. В сжатый файл сохраняются:

- координаты найденного блока
- параметры аффинного преобразования (число от 0 до 7 — преобразование симметрии и сдвиг по яркости для этой пары блоков.)

Для каждого из квадратов перебираем все ранговые области и каждую такую область крутим и пытаемся вписать в доменный блок. Парадокс в том, что сжатый файл не несет никакой информации об исходной файле.

доменные области в ранговые областей Декомпрессия происходит просто преобразования ранговых областей в доменные области

Потери качества:

- подмена области;
- мозаичностью.

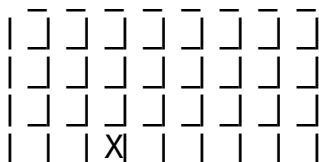
Метод крайне асимметричен. Декомпрессия практически мгновенна.

Метод используется:

- как самостоятельный формат (*.fif); Например в CorelDraw (7, 8, 9). Сжимает лучше (в 10 раз) и дольше по сравнению с jpg. Применяется где требуется 1 раз сжать и много раз воспроизводить.
- промежуточное хранение при редактировании в растровых редакторах. Преобразования происходят не над фотографией а над ее фрактальном представлении.

К следующему занятию придется повторить лекции прошлого семестра:

- IFS
- преобразование Хатчинсона
- Расстояние Хаусдорфа. Посчитать расстояние Хаусдорфа для двух фигур.



		X					

Мат аппарат

Система СИФ:

- T_1 с коэф S_1
- T_2 с коэф S_2
- T_3 с коэф S_3
- T_m с коэф S_m

— множество не пустых компактов \mathbb{R}^2

Вводим преобразование Хатчисона $T(E) = T_1(E) \cap T_2(E) \cap T_3(E) \cap \dots \cap T_m(E)$

E_0 — произв множество $E_1 = T(E_0)$

$E_n = T(E_{n-1})$

$E = \lim_{n \rightarrow \infty} E_n$ в метрики Хаусдорфа.

Существует ряд теорем:

33. Th1

$T_{\text{опр}}(*)$ является сжимающим на пространстве не пустых компактов K с коэффициентом сжатия.

$$S = \max S_1, \dots, S_m$$

34. Th2

Для $\forall E_0 \in K$, СИФ $E_n = T(E_{n-1})$ сх. к ед. $E \in$

$$E = \lim_{n \rightarrow \infty} T^{(n)}(E_0),$$

(по H_0) E — аттрактор СИФ.

Вот мы получим какой-то результат, но на сколько он будет похож на исходный файл.

35. Th3 О коллаже

Ключевая Пусть: I — не \emptyset компактное множество E — аттрактор СИФ. T — преобразование Хатчисона с коэф. S (S из Th1)

Тогда: Если для $\varepsilon > 0$: $H(I, \bigcup_{i=1}^m T_i(I)) < \varepsilon$ То:

$$H(I, E) < \frac{\varepsilon}{1 - S}$$

Эта теорема дает

что при некоторых условиях разница между исходным и полученным изображением поддается оценке сверху.

$H(I, \bigcup_{i=1}^m T_i(I)) < \varepsilon$ — ЭТО ДЛЯ ПЕРВОЙ ИТЕРАЦИИ.

Теорема дает гарантию похожести. Причем на этапе компрессии.

Фрактальное сжатие видео

$$\tau_{m,s,o}(f(x)) = S(x)f(m(x)) + O(x)$$

- $f(x) = f(x, y, t)$ — функция яркости
 - x, y — пространство
 - t — время кадра
- m — смешанная функция
- s — масштабирующая
- O — смещающая

Идея фрактального сжатия видео — она межкадровая. Но эти компрессоры очень непродуцибельны.

Отличия от пространственного метода:

- только 1 итерация
- на требуется сжатие

Вейвлеты (wavelets)

Форматы: *.wi – изображения *.wmv и Indeo 5 AVI — видео

*.wmv —> Window Media 9 и дальше.

Некоторые MPEG4.

Вейвлеты, хоть и не самые эффективные, но они самые оптимальные.

Преобразование Фурье — попытка описать сигнал с помощью не \sin и \cos . Для вейвлетов используется не \sin и \cos а некоторая функция, причем не периодическая.

Волновой анализ — (wavelets — «волночка»)

Основная идея — разложение по базису другой функции.

$$f(x) = \sum_{i=1}^N C_i \Phi_i$$

↓

- $[c_1, \dots, c_N]$ – Канал связи
- N — надо минимизировать

Вейвлеты (wavelets)

Метод сжатия с потерей качества.

Мы уже сказали, что вейвлет преобразование — суть разложение по некоторому базису. Мы будем рассматривать Вейвлет Хаара.

[220.0, 221.0, 212.0, 218.0, 218.0, 217.0, 124.0, 210.0, 209.0]

фильтры:

$$b_i^1 = (a_{2i} + a_{2i+1})/2(\text{низкочастотный фильтр})$$

$$b_i^2 = (a_{2i} - a_{2i+1})/2(\text{высокочастотный фильтр})$$

Результат

$$b^1 = b_1^1 = (215.5, 215.0, 215.5, 209.5)$$

$$b^2 = b_1^2 = (4.5, -3, 1.5, 4.0)$$

фильтры:

$$b_i^1 = (a_{2i} + a_{2i+1})/2(\text{низкочастотный фильтр})$$

$$b_i^2 = (a_{2i} - a_{2i+1})/2(\text{высокочастотный фильтр})$$

Результат

$$b^{12}b_2^1 = (215.25, 212.5)$$

$$b^{12}b_2^2 = (0.25, 3,)$$

фильтры:

$$b_i^1 = (a_{2i} + a_{2i+1})/2(\text{низкочастотный фильтр})$$

$$b_i^2 = (a_{2i} - a_{2i+1})/2(\text{высокочастотный фильтр})$$

Результат

$$b^{111} = b_3^1 = (213.875)$$

$$b^{112} = b_3^2 = (1.375)$$

Чтобы восстановить исходную, нам потребуется знать b_3^1 и b_i^2 .

Пирамида гауссианов

$$b^{111}, b^{11}, b^1, \text{ИСХ}$$

[[[[[[[[[[[+]]]]]]]]]]]

Пирамида лапласианов

$$b^{112} b^{12} b^2 \left[+ \right] \left[+ \right] \left[+ \right] \left[+ \right] \left[+ \right] \left[+ \right] \left[+ \right]$$

Посмотрим 2D

Низкочастотный (по строкам и столбцам — НН):

$$b_{ij}^1 = (a_{2i,2j} + a_{2i+1,2j} + a_{2i,2j+1} + a_{2i+1,2j+1})/4$$

Разность между соседними строками (НВ):

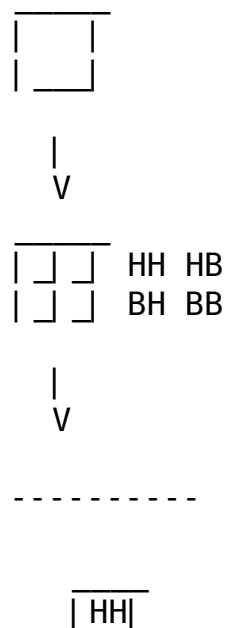
$$b_{ij}^2 = (a_{2i,2j} + a_{2i+1,2j} - a_{2i,2j+1} - a_{2i+1,2j+1})/4$$

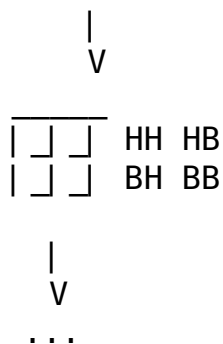
Столбцы (ВН):

$$b_{ij}^3 = (a_{2i,2j} - a_{2i+1,2j} + a_{2i,2j+1} - a_{2i+1,2j+1})/4$$

Диагональ (ВВ):

$$b_{ij}^4 = (a_{2i,2j} - a_{2i+1,2j} - a_{2i,2j+1} + a_{2i+1,2j+1})/4$$





Левый верхний угол — несет уменьшенную копию изображения. Правый нижний обычно вообще выкидывают.

Такое разложение называется пирамидальным разложением.

36. Базис Хаара

$$f(t) = \sum_{k=-\infty}^{\infty} c_k \varphi(t - k) + \sum_{k=-\infty}^{\infty} \sum_{j=0}^{\infty} d_{j,k} \psi(2^j - k)$$

Где c_k и $d_{j,k}$ — коэффициенты

$$\varphi(t) = \begin{cases} 1 & 0 \leq t < 1 \\ 0 & \text{иначе} \end{cases}$$

Вейвлет Хаара:

$$\psi(t) = \begin{cases} 1 & 0 \leq t < 0.5 & \text{низкочастотная} \\ -1 & 0.5 \leq t < 1.0 & \text{высокочастотная} \end{cases}$$

Варианты $\psi(t)$

$\psi(t) =$

- мексиканская шляпа или собреро;
- вейвлет Марле;
- w6 (windows media 9);

Вейвлет-анализом многомасштабе

Требования: Чтобы функция $\psi(t)$ могла наз вейвлетом. Ее среднее значение, интеграл по всей прямой == 0 Чтобы функция быстро убывала к 0 при $t \rightarrow \infty$

Преимущества вейвлет сжатия:

- быстрое
- нет блочных артефактов

Недостатки:

- проблема с пережачей резких граней
- мелких деталей
- появление лестничного эффекта.

Левуович-Маслюк. Дайджест вейвлет анализ.

Eye-Tracker — системы отслеживания движений зрачков глаз

С помощью некоторого оборудования. Система анализирует на какой области человек фокусирует внимание в данный момент. Бывают:

- для управления объектами, людям с нарушением движения — инвалиды
- СНР — системы наведения на цель
- раскладка товара на витринах — маркетинг
- e-маркетинг. веб. анализ юзабилити.
- артефакты сжатия.
- и т.д.

Технические:

- контактные системы
- оптические системы системные дневного света ИК

mai.ru / science / trudy / N 36
usablitylabs.ru

Интернет вещание

37. Файловое

38. Потокное

Ничего про него не знаем. Потокное медиа представляет собой насыщенный интернет-контент, который конечный пользователь, воспринимает как непрерывную последовательность данных в реальном времени.

Применение: Прямые трансляции; Аудио-видео конференции

39. Подготовка к вещанию

Выбор формата вещания. В каком формате на сервер Вид плеера со стороны клиента
Инструмент кодирования

Комплексы софта: MS Windows Media Technology Apple QuickTime Real Media

Кодирование контента. Если идет прямая трансляция.

- 1) Преобразование аналогового контента в цифровую форму.
- 2) Определение скоростей передач.
- 3) Создание файла или потока в формате, который распознается сервером и плеером.
- 4) Передача контента — UDP. CDM — ?
- 5) Воспроизведение.

MS Windows Media Technology: Windows Server 2000 преобразование — Windows Media Encoder формат ASF клиенты подсоединены к Media Service. со стороны клиентов нужен Windows Media Player

ASF — Advanced Streaming Format Внутри — любой кодек. В 1 ASF файле может содержаться несколько копий одного и того же контента, предназначенных для раздачи с разными скоростями. С ним сервер управляется лучше.

У нас 4 варианта раздачи.

40. Виды соединений с клиентом

Виды соединений Unicast — peer2peer — каждый клиент получает свой поток от сервера, индивидуальный Multicast

виды раздачи контента. On-demand Broadcast

Multicast передача от сервера 1 раз. Нагрузка ложится на сетевое оборудование.

- Unicast On-demand — когда клиент индивидуально запрашивает конкретный поток или файл с сервера.

- Unicast Broadcast — поток на сервере один, но раздается он индивидуально — пример web-камеры.
- Multicast On-demand — поток на сервере один, но раздается он всем вместе.

Требования к сетевому оборудованию — оно должно поддерживать IGMP 2.0 (Internet Group Management Protocol) Он должен поддерживаться всем оборудованием.

Второй и третий уровень OSI. Он опирается на:

- Адресация, адреса уровня IP — связывается не с одним клиентом а с группой.
- Динамическая регистрация. Механизм информирования сети о том, что некоторый компьютер является членом некоторой группы рассылки.
- Передача мультикаст пакетов. вопрос установки приоритетов трафика. Учет количества подписчиков.
- Маршрутизация. Построение дерева, которое обеспечивает построения уникального пути следования пакетов.

Портативные чемоданчики. Раздача.

IP-TV мы не касались.

Телевидение

Учебник для вузов — телевидение В.Е.Джаконни
Брайс, справочник по цифровому телевидению.

NTSC — 1953 г США — National Television Standards Committee PAL — западная европа, 1960 — phase-alternating line SECAM — франция, СССР — 1967

Соотношения сторон экранов (телевизоров): 4:3 16:9(10, 11)

кино: 1.85:1 2.35:1 (cinema scope)

Для convert кино 2 телевизор

- обрезание — pan and scan
- letter-boxing — изображение с полосками

41. форма телевизионного экрана

Перевод строки — строчный синхроимпульс и возврат каретки — интервал гашения Форма кадрового синхроимпульса аналогична

42. National Television Standards Committee

США, Канада, ряд стран центральной Америки, Япония, Тайвань.

Полный цветной телевизионный видео сигнал. Содержит в каждой строке составляющую сигнала яркости и сигнала цветности, передаваемую с помощью поднесущей, лежащей в частотной полосе сигнала яркости. поднесущая в каждой строке промодулирована двумя сигналами цветности, чтобы сигналы цветности не создавали взаимных помех, в системе применена квадратурная балансная модуляция.

Характеристики:

- частота развертки — количество полей в секунду 2 поля — 1 кадр.

$$f = 59.940059\text{Гц} = 29.93\text{fps} \approx 30\text{fps}$$

- число видимых строк в кадре — 480.
- высокая цветовая четкость
- декор весьма прост и не содержит линии задержки — низкая стоимость

Недостаток: Очень высокая чувствительность при небольших искажениях в канале цветопередачи.

43. phase-alternating line (PAL)

by Telefunken

Цвет передается с помощью поднесущей.

Через строку производится изменение знака амплитуды поднесущей, это приводит к тому, что цветовая четкость изображения снижена в 2 раза по вертикали. Каждая вторая строка является усреднением предыдущей и последующих строк.

Недостаток:

- цвет фон и насыщенность объектов зависит от уровня освещенности сцены.
- цветовая четкость по вертикали снижена вдвое

44. SECAM

Séquentiel couleur avec mémoire, позднее Séquentiel couleur à mémoire

Используется частотная модуляция поднесущей. Аналогично системе PAL. Поочередно через строку передача цвето-разностных сигналов с дальнейшим восстановлением в декодере путем восстановления строк.

Достоинства:

- цвет фон и насыщенность объектов не зависит от уровня освещенности сцены. Недостаток:
- на резких переходах яркости возникают цветовые окантовки.
- цветовая четкость по вертикали снижена вдвое
- есть линия задержки в 2 строки длиной.

Характеристики:

- 50 Гц — 25 fps
- 625 x 576 видимые строки.

Видеомонтажное оборудование

Трехпостовая система

Принципиальная схема. job/shuttle — непрерывное линейное управление.

45. Интерфейсы передачи видео данных

- Композитный сигнал VHS, Video 8. Он получается путем сложения яркостного и модулированного цветоразностных сигналов, а так же синхроимпульсов. SCART.
- S-video — тюлпан, PS/2, S-VHS, Hi-8
- Y / C Y — яркость C — chrominance.
- Компонентный BetaCam 3 независ YUV + независимый синхроимпульсы – 2
- RGB + Sync (2 штуки) 15 din — VGA-разъем. (+ sync по зеленому)

46. Форматы хранения и записи

что вид сигнала ширина сигнал-шум

VHS композит 12.65 43 дБ

S-VHS Y/C 12.65 45 дБ

Hi8 Y/C 8 44 дБ

BetaCam YUV 12.65 (скорость выше) 48 дБ

Digital BetaCam YUV 12.65 (скорость выше) 51 дБ без сжатия

Dv Dv-Cam DVCPRO YUV 4:1:1 4:2:0 6.35 (скорость выше) 54 дБ 5:1