

1 курс

ПЛАН – КОНСПЕКТ
проведения занятий по дисциплине «Информатика»

Раздел 1. «Информация и информационные процессы»

Тема 1.2: «Информация и ее дискретное представление»

(часть б)

«Цифровое представление аудио и видеоинформации».
Дискретное (цифровое) звуковой информации (ПЗ №10).

Подготовил: преподаватель
В.Н. Борисов

Вопросы занятия:

1. Цифровое представление аудиоинформации. Теоретическая часть. Выполнение практического задания (ПЗ № 10).
2. Цифровое представление видеоинформации.

Время проведения занятий – 4 часа.

Первый вопрос: Цифровое представление аудиоинформации (звуковой информации). Теоретическая часть. Выполнение практического задания (ПЗ № 10).

Использование компьютера для обработки звука началось позднее, нежели чисел, текстов и графики.

Звук – волна с непрерывно изменяющейся амплитудой и частотой. Чем больше амплитуда, тем он громче для человека, чем больше частота, тем выше тон.

Звуковые сигналы в окружающем нас мире необычайно разнообразны. Сложные непрерывные сигналы можно с достаточной точностью представлять в виде суммы некоторого числа простейших синусоидальных колебаний. Причем каждое слагаемое, то есть каждая синусоида, может быть точно задана некоторым набором числовых параметров – амплитуды, фазы и частоты, которые можно рассматривать как код звука в некоторый момент времени.

В процессе кодирования звукового сигнала производится его временная дискретизация– непрерывная волна разбивается на отдельные маленькие временные участки и для каждого такого участка устанавливается определенная величина амплитуды.

Таким образом, непрерывная зависимость амплитуды сигнала от времени заменяется на дискретную последовательность уровней громкости. Каждому уровню громкости присваивается его код. Чем большее количество уровней громкости будет выделено в процессе кодирования, тем большее количество информации будет нести значение каждого уровня и тем более качественным будет звучание.

Качество двоичного кодирования звука определяется глубиной кодирования и частотой дискретизации.

Частота дискретизации – количество измерений уровня сигнала в единицу времени.

Количество уровней громкости определяет глубину кодирования. Современные звуковые карты обеспечивают 16-битную глубину кодирования звука. При этом количество уровней громкости равно $N = 2^{16} = 65536$.

Таким образом, современные звуковые карты обеспечивают кодирование 65 536 уровней сигнала. Каждому значению амплитуды присваивается 16-битный код.

При двоичном кодировании непрерывного звукового сигнала он заменяется последовательностью дискретных уровней сигнала. Качество кодирования зависит от количества измерений уровня сигнала в единицу времени, т.е. частотой дискретизации.

Чем большее количество измерений проводится в 1 секунду (чем больше частота дискретизации), тем точнее процедура двоичного кодирования.

Качество двоичного кодирования звука определяется глубиной кодирования и частотой дискретизации.

Глубина кодирования звука — это количество информации, которое необходимо для кодирования дискретных уровней громкости цифрового звука. Если известна глубина кодирования, то количество уровней громкости цифрового звука можно рассчитывать по общей формуле $N = 2^{16}$. Например, пусть глубина кодирования звука составляет 16 бит, в таком случае количество уровней громкости звука равно: $N = 2^{16} = 65\,536$.

В процессе кодирования каждому уровню громкости звука присваивается свой 16-битовый двоичный код, наименьшему уровню звука будет соответствовать код 0000000000000000, а наибольшему — 1111111111111111.

Частота дискретизации аналогового звукового сигнала может принимать значения от 8 кГц до 48 кГц. При частоте 8 кГц качество дискретизированного звукового сигнала соответствует качеству радиотрансляции, а при частоте 48 кГц — качеству звучания аудио-CD. Следует также учитывать, что возможны как моно-, так и стереорежимы.

Чем больше частота дискретизации и глубина кодирования звука, тем более качественным будет звучание оцифрованного звука и тем лучше можно приблизить оцифрованный звук к оригинальному звучанию.

Второй вопрос: Цифровое представление видеoinформации.

В последнее время компьютер все чаще используется для работы с видеoinформацией. Простейшей такой работой является просмотр кинофильмов и видеоклипов. Следует четко представлять, что обработка видеoinформации требует очень высокого быстродействия компьютерной системы.

Фильм — сочетание звуковой и графической информации. Кроме того, для создания на экране эффекта движения используется дискретная по своей сути технология быстрой смены статических картинок. Исследования показали, что если за одну секунду сменяется более 10-12 кадров, то человеческий глаз воспринимает изменения на них как непрерывные.

Казалось бы, если проблемы кодирования статической графики и звука решены, то сохранить видеоизображение уже не составит труда. Но это только на первый взгляд, поскольку, как показывает разобранный выше пример, 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).

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).

Более универсальным является мультимедийный формат Quick Time, первоначально возникший на компьютерах Apple.

Принципы цифрового представления видео и мультимедийной информации.

Когда говорят о видеозаписи, прежде всего имеют в виду движущееся изображение на экране телевизора или монитора.

Преобразование оптического изображения в последовательность электрических сигналов осуществляется видеокамерой. Эти сигналы несут информацию о яркости и цвете отдельных участков изображения. Они сохраняются на носителе в виде изменения намагниченности видеоленты (аналоговая форма) или в виде последовательности кодовых комбинаций электрических импульсов (цифровая форма).

Процесс превращения непрерывного сигнала в набор кодовых слов называется аналого-цифровым преобразованием.

Это сложный процесс, состоящий из:

- дискретизации, когда непрерывный сигнал заменяется последовательностью мгновенных значений через равные промежутки времени;
- квантования, когда величина каждого отсчёта заменяется округлённым значением ближайшего уровня;
- кодирования, когда каждому значению уровней квантования, полученных на предыдущем этапе, сопоставляются их порядковые номера в двоичном виде.

По своей сути видеофайл — это набор статичных изображений, меняющих друг друга с определенной частотой. Каждое статичное изображение является отдельным кадром видео. Это действительно так, если мы говорим о несжатом видео. Однако в таком формате никто не хранит фильмы.

Дело в том, что несжатое видео занимает на диске очень много места. Кадр видео формата PAL состоит из 720 точек по горизонтали и 576 по вертикали. То есть один кадр состоит из 414720 точек.

Для хранения цвета каждой точки в памяти отводится 24 бита (по 8 бит для каждой из составляющих RGB).

Следовательно, для хранения одного кадра понадобится 9953280 бит (или примерно 1,2 Мбайт).

То есть секунда несжатого видео в формате PAL будет занимать почти 30 Мбайт. А один час такого видео — более 100Гбайт.

Каким же образом полнометражный фильм (а то и несколько) умещается на одном компакт-диске или флеш-накопителе?

Дело в том, что, в основном, видео хранят в видеофайлах, в которых применены различные алгоритмы сжатия информации. Благодаря этим технологиям видеофайл можно сжимать в десятки и сотни раз практически без потери качества картинки и звука.

При работе с цифровым видеосигналом возникает необходимость обработки, передачи и хранения очень больших объемов информации. На современных носителях, таких, как компакт-диск (CD-ROM, 650 Мбайт) или жесткий диск (порядка тысячи мегабайт), сохранить полноценный по времени видеоролик, записанный в поэлементном формате, не удастся. С другой стороны, видеоинформация должна

передаваться со скоростью ее воспроизведения на экране компьютера. Так, полноцветное (24 бит/пиксел) изображение размером 720×576 пикселов из расчета 25 кадр/с требует скорости передачи видеоданных 240 Мбит/с. Однако пропускная способность каналов ЛВС FDDI — порядка 100-200 Мбит/с, а Ethernet — всего 10 Мбит/с. Поэтому использование видеоданных в составе электронных изданий оказывается невозможным.

Развитие технологий перевода видеоинформации в цифровой формат и их дальнейшее применение в цифровом ТВ поставили проблему сжатия видеоданных в ряд наиболее важных. Ее положительное решение оказалось возможным лишь на базе разработки эффективных методов и алгоритмов сжатия видеоданных.

Следует отметить, что традиционные алгоритмы сжатия данных без потерь здесь практически неприменимы, поскольку дают для реальной видеоинформации слишком незначительный выигрыш. Например, алгоритмы, основанные на компрессии за счет кодирования длинами серий и адресно-позиционного кодирования (RLE, LZ, LZW и т.п.), не дают должного эффекта.

Для покадрового сжатия видеоданных можно использовать алгоритмы компрессии статической графики — сжатие с потерями (JPEG). При этом восстановленное изображение кадра, как правило, не совпадает с оригиналом. Однако реализация таких алгоритмов достаточно сложна и процессы декодирования требуют значительных затрат времени. Видеоинформация накладывает специфические ограничения на скорость декодирования данных: декодер (аппаратно-программное средство, осуществляющее декомпрессию данных) должен успевать разархивировать изображение за $1/25$ с, пока на экране отображается предыдущий кадр. Данное ограничение не дает возможности реализовать алгоритмы с большей степенью сжатия.

Еще одно ограничение — сложность аппаратной реализации декодирующих устройств. В реальных приложениях (цифровые видеокамеры, видеотелефоны, видеофоны и т.д.) оптимальным решением проблемы оказывается реализация алгоритма на заказном наборе микросхем с ограниченным числом транзисторов в чипе. Поэтому реализация подобных быстродействующих декодирующих аппаратно-программных устройств не всегда возможна.

Реальным решением проблемы стало сжатие всего видеоряда, включающего последовательность видеок кадров.

Стандартным методом цифрового кодирования на компьютере является РСМ (Pulse Code Modulation). Наиболее популярным форматом, используемым для хранения несжатых аудиоданных, является Microsoft РСМ (WAV). Для видеороликов стандартным для компьютера считается Microsoft Audio/Video Interleaved (AVI). Сжатие аудио- или видеоданных как процесс подразумевает конвертацию соответственно несжатого WAV- или AVI- файла в другой формат с использованием алгоритма сжатия (поэтому программы для компрессии/декомпрессии данных называют конверторами). При этом может быть использован любой формат (даже WAV и AVI), если он поддерживает этот алгоритм.

Важную роль в решении проблемы сжатия видеоданных сыграли результаты, полученные группой комитета по стандартизации MPEG (Motion Pictures Experts Group). Эта группа предложила технологию компактного представления цифровых видео- и аудиосигналов. Основная идея заключалась в преобразовании потока

дискретных цифровых данных в поток некоторых записей, которые требовали меньшего объема памяти. Это преобразование основано на использовании статистической избыточности и особенностях человеческого восприятия. Закодированные независимо аудио- и видеопотоки в дальнейшем связываются системным потоком, который осуществляет синхронизацию и объединение множества потоков различных данных в одну кодовую последовательность.

Разработанный этой группой метод сжатия и соответствующие форматы семейства MPEG унаследовали многое в своей структуре от JPEG. Однако противоположность графическим форматам MPEG использовал кодирование отличий последующих кадров от некоторых опорных изображений кадров. В 1990 г. был создан формат MPEG-1, который ориентировался на сжатие видео- и аудиоинформации.

После разработки первого стандарта на сжатие видео- и аудиоинформации эта же группа создала формат и соответствующую технологию, ориентированную на применение к видеоданным с более высоким разрешением и потокам, в том числе для эффективного представления видеоинформации вещательного качества (SDTV — Standard Definition Television). Эффективное представление чересстрочной развертки (вещательного) видеосигнала оказалось более сложной задачей, чем прямые (не чересстрочные) сигналы, определенные MPEG-1. MPEG-2 ввел схему декорреляции многоканального дискретного аудиосигнала объемного звука, используя более высокий фактор избыточности.

В дальнейшем были созданы форматы MPEG-3, MPEG-4, MPEG-7, MPEG-J.

Сегодня MPEG — единый формат представления данных спецификации United States Grand Alliance HDTV, группы European Digital Video Broadcasting и Digital Versital Disc (DVD).

В литературе MPEG может подразделяться на фазы (MPEG-1, MPEG -2, MPEG-4 и т.д.), а в области аудиоинформации — еще и на уровни (layers). Фазы обозначаются арабскими цифрами, уровни — римскими. Некоторые фазы MPEG так и не были закончены. Например, разработка MPEG-3, предназначенного для телевидения высокой четкости (HDTV) с размерами кадров 1920×1080 при частоте смены 30 кадр/с и силой сжатия 20-40 Мбит/с, не была завершена, поскольку оказалось, что эта область поддерживается форматом MPEG -2. Нет информации о разработке MPEG -6, который предназначался для беспроводной передачи данных; MPEG -8, цель которого — четырехмерное описание объектов.

Экспертная группа по мультимедиа и гипермедиа MHEG (Multimedia Hypermedia Expert Group) определила стандарт для обмена мультимедийными объектами (видео, звук, текст и другие произвольные данные) между приложениями и передачи их разными способами (локальная сеть, сети телекоммуникаций и вещания) с использованием MHEG object classes. Этот стандарт позволил программным объектам включать в себя любую систему кодирования (например, MPEG), которая определена в базовом приложении. MHEG был принят советом по цифровому видео и звуку (DAVIC — Digital Audio-Visual Council). MHEG-объекты создаются мультимедийными приложениями.

Считается, что MHEG — будущий международный стандарт для интерактивного TV, так как он работает на любых платформах и его документация свободно распространяема.

Наряду с указанными стандартными форматами имеется конечное множество форматов кодирования видео- и аудиоинформации, предложенных фирмами, производящими различные программные приложения. К ним можно отнести: формат RealAudio, разработанный фирмой RealNetworks, для хранения сжатых голосовых аудиоданных (речи); формат аудиоданных SoundVQ, разработанный компанией Yamaha; формат Windows Media Technology 4.0, представленный фирмой Microsoft, поддерживает потоковую передачу данных в Internet и имеет продвинутую систему сжатия аудио- и видеоданных; формат QuickTime фирмы Apple был разработан для использования в мультимедийных приложениях на компьютерах Macintosh и т.п.