

Базовая информационная технология: мультимедиа-технологии

1. Понятие мультимедиа

Термин «мультимедиа» с английского можно перевести как «многие среды» (от multi – много и media – среда). Это комплекс аппаратных и программных средств, позволяющих пользователю работать в диалоговом режиме с мультимедийными данными, организованными в единой информационной среде.

В настоящее время мультимедиа-технологии являются бурно развивающейся областью информационных технологий. В этом направлении активно работает значительное число крупных и мелких фирм, технических университетов и студий. Области применения чрезвычайно многообразны: интерактивные обучающие и информационные системы, САПР и др.

Основными характерными **особенностями** этих технологий являются:

- объединение многокомпонентной информационной среды (текста, звука, графики, фото, видео) в однородном цифровом представлении;
- обеспечение надежного (отсутствие искажений при копировании) и долговечного хранения (гарантийный срок хранения – десятки лет) больших объемов информации;
- простота переработки информации.

Достигнутый технологический базис основан на использовании нового стандарта оптического носителя DVD (Digital Versatile/Video Disk), имеющего емкость порядка единиц и десятков гигабайт и заменяющего все предыдущие: CD-ROM, Video-CD, CD-audio. Использование DVD позволило реализовать концепцию однородности цифровой информации и повысить качество звука и изображения. Одно устройство заменяет аудиоплеер, видеомаягнитофон, CD-ROM, дисковод, слайдер и др.

2. Классификация мультимедиа

Мультимедиа может быть классифицирована на две основные группы:

- **Линейная.** Аналогом линейного способа представления является кино. Человек, просматривающий данный документ, никаким образом не может повлиять на его вывод.
- **Нелинейная.** Нелинейный способ представления информации позволяет человеку участвовать в выводе информации, взаимодействуя каким-либо образом со средством отображения мультимедийных данных. Участие человека в данном процессе также называется интерактивностью.

Нелинейный способ представления мультимедийных данных иногда называется термином «гипермедиа».

В качестве примера линейного и нелинейного мультимедийных продуктов, можно рассматривать проведение компьютерных презентаций. Компьютерная презентация представляет собой последовательность слайдов, содержащих мультимедийные объекты. Переход между слайдами осуществляется с помощью управляющих объектов (кнопок) или гиперссылок. В некоторых случаях презентацию запускают в автоматическом режиме, и она повествует о чем-то без участия человека (например, проведение различных выставок). Этот способ донесения информации может быть назван линейным. В случае же живой презентации аудитория имеет возможность взаимодействовать с докладчиком (например, задавать ему вопросы), что позволяет ему отходить от темы презентации, поясняя некоторые термины или более подробно освещая спорные части доклада. Таким образом, живая презентация может быть представлена, как нелинейный (интерактивный) способ подачи информации.

3. Структурные компоненты мультимедиа

Многокомпонентную мультимедиа-среду разделяют на группы: текст, аудиоряд, видеоряд и графическая информация.

3.1 Текст

Текст – это упорядоченный набор предложений, связанных по смыслу. В смысловой цельности текста отражаются те связи и зависимости, которые имеются в самой действительности (общественные события, явления природы, человек, его внешний облик и внутренний мир, предметы неживой природы и т.д.).

Восприятие текста изучается в рамках таких дисциплин, как лингвистика текста и психолингвистика.

Текстовый файл – обычная форма представления текста на компьютере. Каждый символ из используемого набора символов кодируется в виде одного байта, а иногда в виде последовательности подряд идущих двух, трех и более байтов.

Особой разновидностью текстовых данных следует считать т.н. гипертекст. Термин «гипертекст» был введен Тедом Нельсоном в 1965 году для обозначения «текста ветвящегося или выполняющего действия по запросу». Обычно гипертекст представляется набором текстов, содержащих узлы перехода от одного текста к какому-либо другому, позволяющие избирать читаемые сведения или последовательность чтения. Общеизвестным и притом ярко выраженным примером гипертекста служат веб-страницы – документы на HTML (гипертекстовом языке разметки), размещенные в интернете.

Существуют стилистические, жанровые и тематические классификации текста.

3.2 Аудио

Аудио (от лат. audio – «слышу») – общий термин, относящийся к звуковым технологиям. Как правило, под термином аудио понимают звук, записанный на звуковом носителе, а также запись и воспроизведение звука, звукозаписывающая и звуковоспроизводящая аппаратура.

Таким образом, аудиальный компонент мультимедийного продукта предназначен для передачи звуковых данных. Как физическое явление звук изучается в рамках акустики, но при этом акустика является междисциплинарной наукой, использующей для решения своих проблем широкий круг дисциплин: математику, физику, психологию, архитектуру, электронику, биологию, теорию музыки и др. Непосредственное отношение к вопросам мультимедиа-технологий имеют такие направления современной акустики, как музыкальная акустика, электроакустика, акустика речи, цифровая акустика.

Аудиоряд мультимедиа-среды может включать речь, музыку и звуковые эффекты (звуки типа грома, шума, скрипа и т. д.).

Звук представляет собой звуковую волну с непрерывно меняющейся амплитудой и частотой. Чем больше амплитуда сигнала, тем он громче для человека, чем больше частота сигнала, тем выше тон.

Музыкальный звук обладает следующими характеристиками:

- определенной высотой (обычно от 16 до 4500 Гц);
- тембром, который определяется присутствием в звуке обертонов и зависит от источника звука;
- громкостью, которая не может превышать болевого порога;
- длительностью.

Речевой звук образуется произносительным аппаратом человека с целью языкового общения. Звуки речи подразделяются на шумы и тоны. Тоны в речи возникают в результате колебания голосовых связок; шумы образуются вследствие неперiodических колебаний выходящей из легких струи воздуха. Период основного тона разных людей (мужчин, женщин, детей) находится в диапазоне 50-250 Гц.

Среди звуковых носителей информации выделяют аналоговые и цифровые носители. Примером аналогового хранения звуковой информации является виниловая пластинка, а дискретного – аудиокомпакт-диск. Для того чтобы компьютер мог обрабатывать звук, непрерывный звуковой сигнал должен быть превращен в последовательность электрических импульсов (двоичных нулей и единиц). Таким образом, непрерывная зависимость амплитуды сигнала от времени $A(t)$ заменяется на дискретную последовательность уровней громкости, т.

е. производится *временная дискретизация*. На графике это выглядит как замена гладкой кривой на последовательность «ступенек» – рис. 1.

Каждой «ступеньке» присваивается значение уровня громкости звука, его код (1, 2, 3 и так далее). Уровни громкости звука можно рассматривать как набор возможных состояний, соответственно, 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.

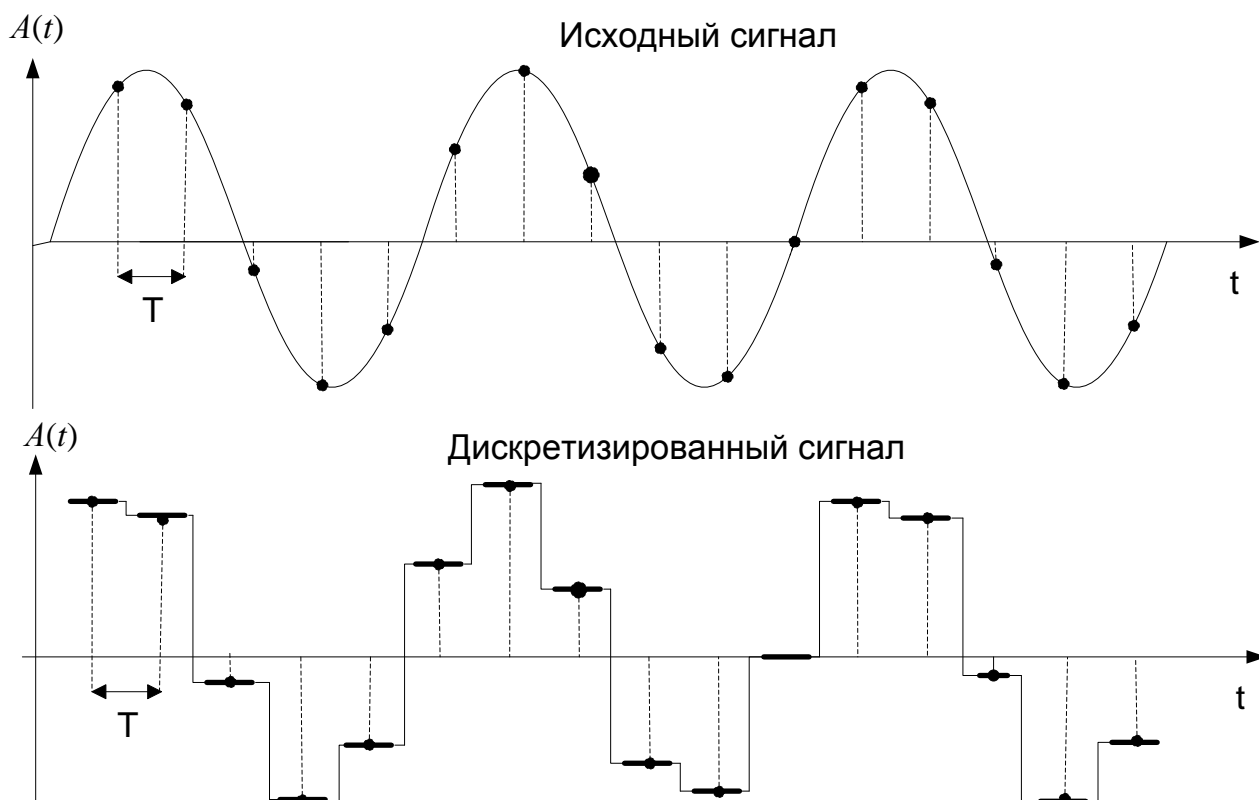


Рис. 1. Перевод непрерывной информации в дискретную

Современные звуковые карты обеспечивают 16-битную глубину кодирования звука. Количество различных уровней сигнала (состояний при данном кодировании) можно рассчитать по формуле:

$$N = 2^I = 2^{16} = 65536,$$

где I — глубина звука.

Таким образом, современные звуковые карты могут обеспечить кодирование 65536 уровней сигнала. Каждому значению амплитуды звукового сигнала присваивается 16-битный код.

Качество двоичного кодирования звука определяется *глубиной кодирования* и *частотой дискретизации*.

Количество измерений в секунду может лежать в диапазоне от 8000 до 48 000, то есть частота дискретизации аналогового звукового сигнала может принимать значения от 8 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 стерео-режимы.

Можно оценить информационный объем стереоаудиофайла длительностью звучания 1 секунда при высоком качестве звука (16 битов, 48 кГц). Для этого количество битов, приходящихся на одну выборку, необходимо умножить на количество выборок в 1 секунду и умножить на 2 (стерео): $16 \text{ бит} \cdot 48\,000 \cdot 2 = 1\,536\,000 \text{ бит} = 192\,000 \text{ байт} = 187,5 \text{ Кбайт}$. Для записи одной минуты WAV-звука высшего качества необходима память порядка 10 Мбайт (**$187,5 \text{ Кбайт} \cdot 60 = 10 \text{ Мбайт}$**), поэтому стандартный объем CD (до 640 Мбайт) позволяет записать не более часа (**$10 \text{ Мбайт} \cdot 60 = 600 \text{ Мбайт}$**) WAV.

В классификации форматов аудио-файлов выделяют форматы без потерь и форматы с потерями. Аудиоформаты без потерь предназначены для точного (с точности до частоты дискретизации) представления звука. В свою очередь они делятся на несжатые и сжатые форматы.

Примеры несжатых форматов:

- RAW – сырые замеры без какого-либо заголовка или синхронизации.

- WAV (Waveform audio format) – разработан Microsoft совместно с IBM, распространенная форма представления звуковых данных небольшой продолжительности.

- CDDA – стандарт для аудио-CD. Первая редакция стандарта издана в июне 1980 года компаниями Philips и Sony, затем была доработана организацией Digital Audio Disc Committee.

Примеры сжатых форматов:

- WMA (Windows Media Audio 9 Lossless) – лицензируемый формат аудио-файлов, разработанный компанией Microsoft для хранения и трансляции. В рамках формата есть возможность кодирования звука как с потерей, так и без потери качества.

- FLAC (Free Audio Lossless Audio Codec) – популярный формат для сжатия аудиоданных. Поддерживается многими аудио-приложениями, а также устройствами воспроизведения звука.

Аудиоформаты с потерями ориентированы в первую очередь на компактное хранение звуковых данных: при этом идеально точное воспроизведение записанного звука не гарантируется. Примеры таких форматов:

- MP3 – лицензируемый формат файла для хранения аудиоинформации, разработанный рабочей группой института Фраунхофера MPEG в 1994 году. На данный момент MP3 является самым известным и популярным из распространенных форматов цифрового кодирования звуковой информации с потерями. Он широко используется в файлообменных сетях для передачи музыкальных произведений. Формат может проигрываться в любой современной операционной системе, на

практически любом портативном аудио-плеере, а также поддерживается всеми современными моделями музыкальных центров и DVD-плееров.

- Vorbis – свободный формат сжатия звука с потерями, появившийся летом 2002 года. Психоакустическая модель, используемая в Vorbis, по принципам действия близка к MP3. По всевозможным оценкам этот формат является вторым по популярности после MP3 форматом компрессии звука с потерями. Широко используется в компьютерных играх и в файлообменных сетях для передачи музыкальных произведений.

- AAC (Advanced Audio Coding) – формат аудио-файла с меньшей потерей качества при кодировании, чем MP3 при одинаковых размерах. Изначально создавался как преемник MP3 с улучшенным качеством кодирования, но в настоящий момент распространен существенно меньше, чем MP3.

3.3 Видео

Видео (от лат. video – «смотрю», «вижу») – под этим термином понимают широкий спектр технологий записи, обработки, передачи, хранения и воспроизведения визуального и аудиовизуального материала на мониторах.

Видеоряд по сравнению с аудиорядом характеризуется большим числом элементов. Выделяют статический и динамический видеоряды.

Статический видеоряд включает компьютерную графику (рисунки, интерьеры, поверхности, символы в графическом режиме) и фото (фотографии и сканированные изображения).

Динамический видеоряд представляет собой последовательность статических элементов (кадров). Можно выделить три типовых группы:

- квазивидео – разреженная последовательность фотографий (6–12 кадров в секунду);
- обычное видео (life video) – последовательность фотографий (около 24 кадров в секунду);
- анимация – последовательность рисованных изображений.

Пространственная дискретизация

В процессе кодирования изображения производится его пространственная дискретизация. Изображение разбивается на отдельные маленькие фрагменты (точки), причем каждому фрагменту присваивается значение его цвета, т.е. код цвета (рис. 2).

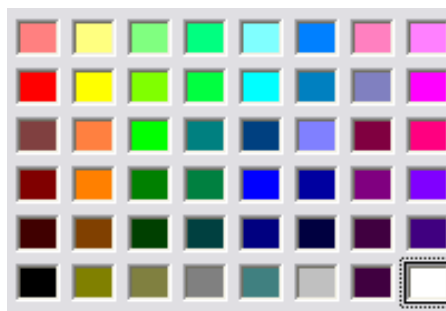


Рис. 2. Пространственная дискретизация изображения

Наиболее важные характеристики видеосигнала – это количество кадров в секунду, развертка, разрешение, соотношение сторон, цветовое разрешение, ширина видеопотока, качество. Рассмотрим эти характеристики по отдельности.

Количество кадров в секунду (частота) – это число неподвижных изображений, сменяющих друг друга при показе 1 секунды видеоматериала и создающих эффект движения на экране. Чем больше частота кадров, тем более плавным и естественным будет казаться движение. Минимальный показатель, при котором движение будет восприниматься однородным – примерно 10 кадров в секунду (это значение индивидуально для каждого человека). Компьютерные оцифрованные видеоматериалы хорошего качества, как правило, используют частоту 30 кадров в секунду.

Развертка видеоматериала может быть прогрессивной (построчной) или чересстрочной (интерлейсинг). При прогрессивной развертке все горизонтальные линии (строки) изображения отображаются одновременно, при чересстрочной – показываются попеременно четные и нечетные строки. Чересстрочная развертка была изобретена для показа изображения на кинескопах и используется сейчас для передачи видео по «узким» каналам, не позволяющим передавать изображение во всем качестве.

Разрешающая способность экрана и глубина цвета. Любой видеосигнал характеризуется вертикальным и горизонтальным разрешением, измеряемым в пикселах. Обычное аналоговое телевизионное разрешение составляет 720×576 пикселей. Новый стандарт высокоотчетливого цифрового телевидения HDTV предполагает разрешения до 1920(количество точек в каждой строке)×1080 (количество строк) с прогрессивной разверткой.

Пример, рассмотрим формирование на экране монитора растрового изображения, состоящего из 600 строк по 800 точек в каждой строке (всего 480 000 точек). В простейшем случае (черно-белое изображение без градаций серого цвета) каждая точка экрана может иметь одно из двух состояний – «черная» или «белая», т. е. для хранения ее состояния необходим 1 бит. ($N = 2 = 2^I$, $I = 1$ бит, где I – глубина цвета, N – количество цветов).

Цветные изображения формируются в соответствии с двоичным кодом цвета каждой точки, хранящимся в видеопамяти. Цветные изображения могут иметь различную глубину цвета, которая задается количеством битов, используемым для кодирования цвета точки. Наиболее распространенными значениями глубины цвета являются $I = 8, 16, 24$ или 32 бита (таблица 1).

Качество двоичного кодирования изображения определяется *разрешающей способностью экрана и глубиной цвета*.

Количество цветов и цветовое разрешение видеосигнала описывается цветовыми моделями. В компьютерной технике применяется в основном RGB и HSV.

Цветовая модель называется RGB -моделью по первым буквам английских названий цветов (Red, Green, Blue). Цветное изображение на экране монитора формируется за счет смешивания трех базовых цветов: красного, зеленого и синего.

Таблица 1.

Формирование цветов при глубине цвета 24 бита

| Название цвета | Интенсивность | | |
|----------------|---------------|----------|----------|
| | Красный | Зеленый | Синий |
| Черный | 00000000 | 00000000 | 00000000 |
| Красный | 11111111 | 00000000 | 00000000 |
| Зеленый | 00000000 | 11111111 | 00000000 |
| Синий | 00000000 | 00000000 | 11111111 |
| Голубой | 00000000 | 11111111 | 11111111 |
| Желтый | 11111111 | 11111111 | 00000000 |
| Белый | 11111111 | 11111111 | 11111111 |

Для того чтобы на экране монитора формировалось изображение, информация о каждой его точке (код цвета точки) должна храниться в видеопамяти компьютера. Рассчитаем необходимый объем видеопамяти для одного из графических режимов, например, с разрешением 800 x 600 точек и глубиной цвета 24 бита на точку.

Всего точек на экране: $800 \cdot 600 = 480\,000$.

Необходимый объем видеопамяти:

$24 \text{ бит} \cdot 480\,000 = 11\,520\,000 \text{ бит} = 1\,440\,000 \text{ байт} = 1406,25 \text{ Кбайт} = 1,37 \text{ Мбайт}$.

Соотношение ширины и высоты кадра – важнейший параметр в любом видеоматериале. Старому стандарту, который предписывает соотношение сторон как 4:3, появившемуся еще в 1910 году, на смену приходит более соответствующий естественному полю зрения человека стандарт 16:9, на который сейчас ориентируется цифровое телевидение.

Ширина видеопотока или битрейт (от англ. bit rate – частота битов) – это количество обрабатываемых бит видеоинформации за секунду времени. Чем выше ширина видеопотока, тем в общем лучше качество видео. Например, для формата VideoCD битрейт составляет всего примерно 1 Мбит/с, для DVD – около 5 Мбит/с, а для формата HDTV – около 10 Мбит/с.

Качество видео измеряется с помощью формальных метрик, таких, как PSNR или SSIM, или с использованием субъективного сравнения с привлечением экспертов.

Из современных стандартов цифрового кодирования и сжатия видеоматериалов можно выделить следующие:

- MPEG-2 – группа стандартов цифрового кодирования видео и аудио сигналов. MPEG-2 в основном используется для кодирования видео и аудио при вещании, включая спутниковое вещание и кабельное телевидение. С некоторыми модификациями этот формат также используется как стандарт для сжатия DVD.

- MPEG-4 – новый международный стандарт сжатия цифрового видео и аудио, появившийся в 1998 году. Используется для вещания (потокное видео), записи дисков с фильмами, видеотелефонии и широковещения. Включает в себя многие функции MPEG-2 и других стандартов, добавляя такие функции, как поддержка языка виртуальной разметки VRML для показа 3D-объектов, объектно-ориентированные файлы, поддержка управления правами и разные типы интерактивного медиа.

- Ogg Theora – видеокодек, разработанный Фондом Xiph.Org как часть их проекта «Ogg» (целью этого проекта является интеграция видеокодека On2 VP3, аудиокодека Ogg Vorbis и мультимедиа-контейнера Ogg в одно мультимедийное решение, наподобие MPEG-4). Полностью открытый, свободный в лицензионном отношении мультимедиа-формат.

3.4 Компьютерная графика

Данное направление мультимедийных технологий предназначено для передачи пользователю визуальных изображений. Первые вычислительные машины не имели отдельных средств работы с графикой, однако уже использовались для получения и обработки изображений. Существенный прогресс компьютерная графика испытала с появлением возможности запоминать изображения и выводить их на компьютерном дисплее.

Для передачи и хранения цвета в компьютерной графике используются различные формы его представления. В общем случае цвет

представляет собой набор чисел, координат в некоторой цветовой системе. Известны, например, следующие модели цветопередачи:

- RGB (аббревиатура английских слов Red, Green, Blue – красный, зеленый, синий) – аддитивная цветовая модель: цветное изображение на экране монитора формируется за счет смешивания трех базовых цветов: красного, зеленого и синего. Иначе говоря, если цвет экрана, освещенного цветным прожектором, обозначается как (r_1, g_1, b_1) , а цвет того же экрана, освещенного другим прожектором, – (r_2, g_2, b_2) , то при освещении двумя этими прожекторами цвет экрана будет обозначаться как $(r_1+r_2, g_1+g_2, b_1+b_2)$. Выбор основных цветов обусловлен особенностями физиологии восприятия цвета сетчаткой человеческого глаза. Цветовая модель RGB нашла широкое применение в технике. В телевизорах и мониторах применяются три электронные пушки (либо три вида светодиодов, светофильтров и др.) для красного, зеленого и синего каналов.

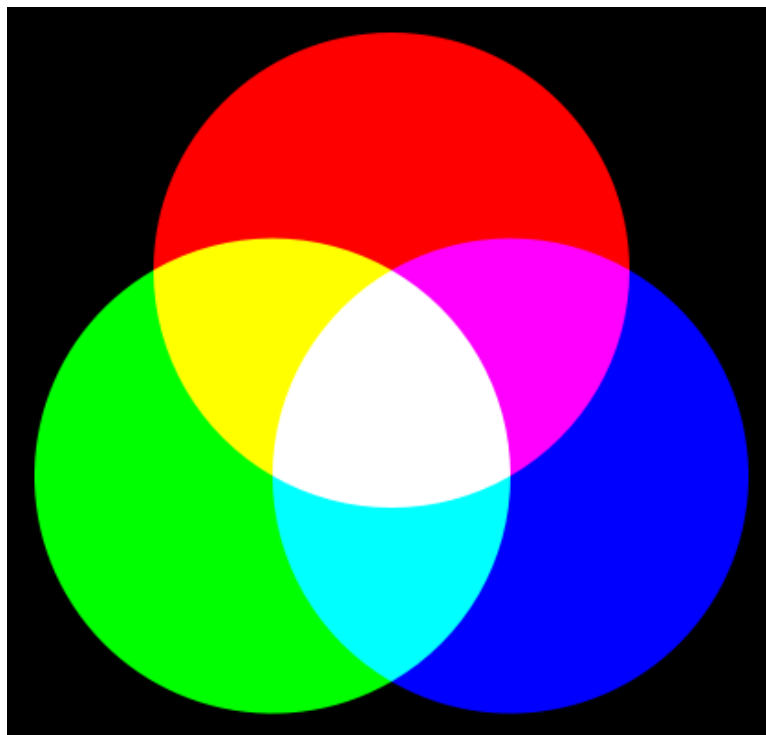


Рис. 3. Система цветопередачи RGB

- CMYK (от англ. Cyan, Magenta, Yellow, black – голубой, пурпурный, желтый, черный) – субтрактивная схема формирования цвета, используемая обычно в полиграфии для стандартной триадной печати.

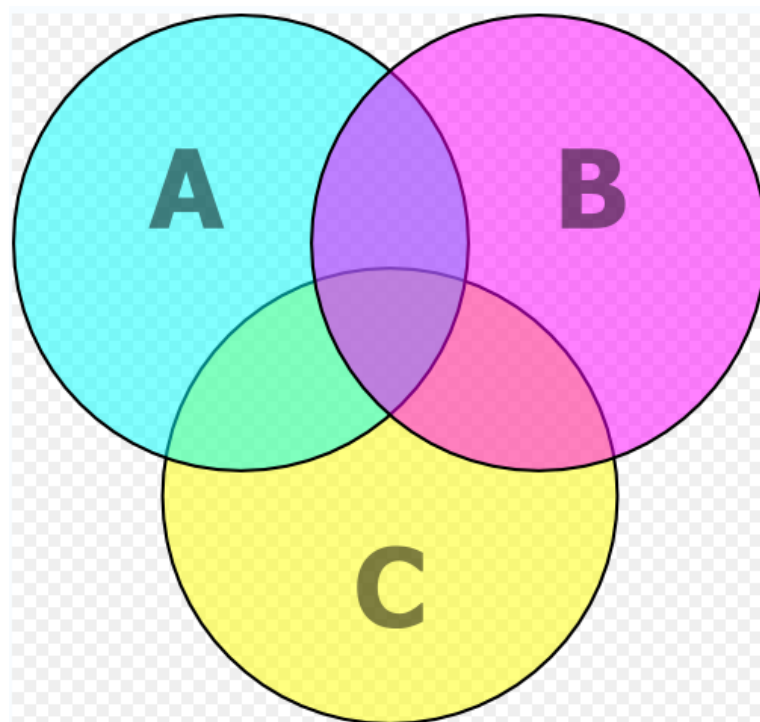


Рис. 4. Система цветопередачи CMYK

- HSV (от англ. Hue, Saturation, Value – тон, насыщенность, значение) – цветовая модель, в которой координатами являются цветовой тон, насыщенность (называемая также чистотой цвета) и значением (яркостью) цвета. Данная модель является нелинейным преобразованием модели RGB.

Двумерная и трехмерная графика

По способам построения изображений компьютерную графику можно разделить на двумерную и трехмерную графику. Двумерная компьютерная графика (2D) классифицируется по типу представления графической информации, и следующими из него алгоритмами обработки изображений. Известны следующие виды двумерной графики:

- Растровая графика. Эта разновидность двумерной графики всегда оперирует двумерным массивом (матрицей) пикселей. Пиксел (или пиксель) – мельчайшая единица растрового изображения, представляющая собой неделимый объект прямоугольной (обычно квадратной) формы, обладающий определенным цветом. Без особых потерь визуального качества растровые изображения можно только уменьшать; увеличение же растровых изображений приводит к увеличению дискретности изображения (см. рис. 5). В растровом виде представимо любое изображение, однако этот способ хранения характеризуется большим объемом памяти, необходимым для работы с изображениями и потерями при редактировании.

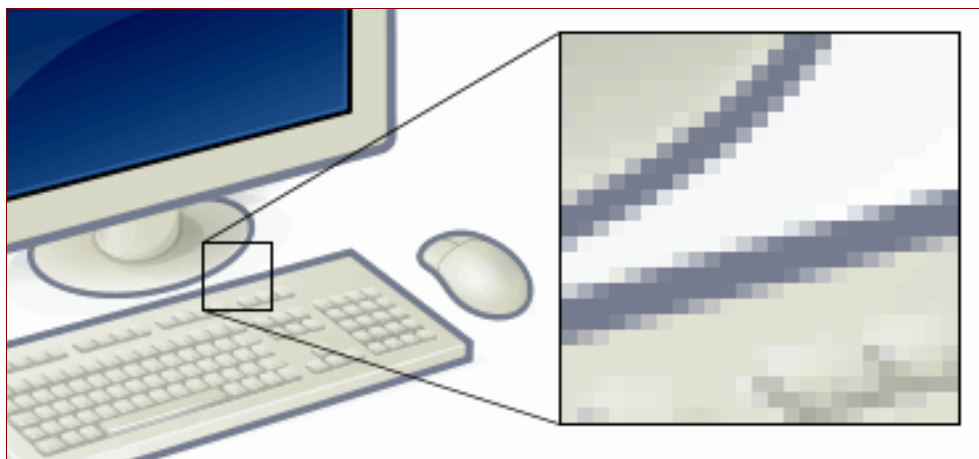


Рис. 5. Результат увеличения растрового изображения

• Векторная графика. Представляет изображение как набор примитивов, в качестве которых обычно выбираются точки, прямые, окружности, прямоугольники, а также сплайны некоторого порядка. Объектам присваиваются некоторые атрибуты (толщина линий, цвет заполнения и т.д.). Рисунок хранится как набор координат, векторов и других численных значений, характеризующих набор примитивов. Изображение в векторном формате дает простор для редактирования, поскольку может без потерь (в отличие от растрового изображения) масштабироваться, поворачиваться, деформироваться. Вместе с тем, не всякое изображение может быть представлено в виде набора примитивов. Такой способ представления хорош для схем, используется для масштабируемых шрифтов, деловой графики, очень широко используется для создания мультфильмов и просто роликов разного содержания.



Рис. 6. Векторные изображения для мебели и матирования стекол в формате CDR(CorelDRAW)



• Фрактальная графика. Фракталов в общем смысле называется объект, отдельные элементы которого наследуют свойства родительских структур. Поскольку более детальное описание элементов меньшего масштаба происходит по простому алгоритму, описать такой объект можно всего лишь несколькими математическими уравнениями.

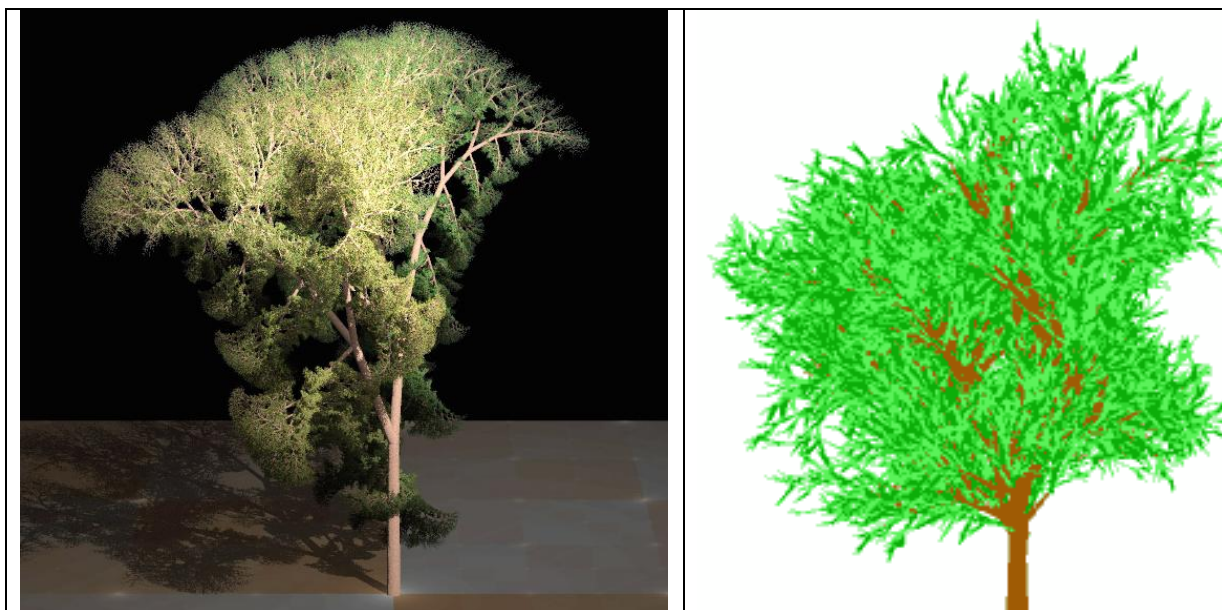


Рис. 7. Фрактальные деревья

Фракталы широко применяются в компьютерной графике для построения изображений природных объектов, таких, как деревья, кусты, горные ландшафты, поверхности морей и т. д.

- Трехмерная компьютерная графика (3D) оперирует с объектами в трехмерном пространстве. Обычно результаты визуализации трехмерной графики представляют собой плоскую картинку, проекцию. В трехмерной графике все объекты обычно представляются как набор поверхностей или частиц. Минимальную поверхность называют полигоном. В качестве полигона чаще всего выбирают треугольники.

Компьютерная графика представляет собой одно из наиболее мощных современных направлений развития компьютерных технологий.

4. Применение мультимедиа-технологий

Мультимедиа-технологии нашли широкое применение в таких сферах человеческой деятельности, как искусство, образование, индустрия развлечений, медицина, бизнес, научные исследования и др. В настоящее время мультимедийный способ передачи информации стал неотъемлемым элементом современных компьютерных систем.

Основные направления использования мультимедиа-технологий:

- электронные издания для целей образования и др.;
- в телекоммуникациях со спектром возможных применений от просмотра заказной телепередачи и выбора нужной книги до участия в мультимедиа-конференциях. Такие разработки получили название Information Highway;
- мультимедийные информационные системы («мультимедиа-киоски»), выдающие по запросу пользователя наглядную информацию.

5. Технические и программные средства мультимедиа-технологии

Для построения мультимедиа системы необходима дополнительная аппаратная поддержка: аналого-цифровые и цифроаналоговые преобразователи для перевода аналоговых аудио- и видеосигналов в цифровой эквивалент и обратно, видеопроцессоры для преобразования обычных телевизионных сигналов к виду, воспроизводимому электронно-лучевой трубкой дисплея, декодеры для взаимного преобразования телевизионных стандартов, специальные интегральные схемы для сжатия данных в файлы допустимых размеров и т. д.

Аппаратные средства.

Все оборудование, отвечающее за звук, объединяется в звуковые карты, а за видео – в видеокарты.

Аппаратные средства мультимедиа:

- Средства звукозаписи (звуковые платы, микрофоны);

- Средства звуковоспроизведения (усилитель, колонки, акустические системы, наушники и гарнитуры);
- Манипуляторы (компьютерные мыши, джойстики, миди-клавиатуры);
- Средства «виртуальной реальности» (перчатки, очки, шлемы виртуальной реальности, используемые в играх);
- Носители информации (CD, DVD и HDD);
- Средства передачи (мини видеокамеры, цифровые фотоаппараты);
- Средства записи (приводы CD / DVD-ROM , CDRW / DVD+RW, TV- и FM-тюнеры);
- Средства обработки изображения (платы видеомонтажа, клавиатуры, графические акселераторы).
- Компьютер, телевизор, средства для получения и удобного восприятия информации и др.

Программные средства мультимедиа технологий.

Программные средства мультимедиа складываются из трех компонентов:

1. Системные программные средства.
2. Инструментальные программные средства.
3. Прикладные программные средства.

Системные программные средства – это набор программ, входящих в состав операционной системы компьютера и осуществляющих управление устройствами мультимедиа, причем это управление на двух уровнях – физическое управление вводом-выводом информации на низком уровне с помощью машинных команд и управление пользователем характеристиками устройств с помощью графического интерфейса, изображающего пульт управления устройством, например регулировки громкости звука, тембра, стереобаланса и т. д. Как правило, программы физического управления устройствами называют драйверами устройств.

Инструментальные программные средства – программы позволяющие модифицировать мультимедийные файлы и создавать мультимедийные приложения.

Инструментальные программные средства – это пакеты программ для создания мультимедийных приложений:

- редакторы неподвижных графических изображений,
- средства создания анимированных GIF-файлов,
- средства аудио- и видеомонтажа,
- средства создания презентаций,
- средства распознавания текстов, введенных со сканера,
- средства создания обучающих программ,
- системы распознавания голоса и преобразования звуковых файлов в текстовые,
- системы создания приложений виртуальной реальности и другие.

Инструментальные средства существенно расширяют возможности управления мультимедийными устройствами по сравнению с теми, которые предоставляют системные средства, но это всегда платные продукты и некоторые из них стоят очень дорого, например профессиональные системы видеомонтажа.

Прикладные программные средства – это готовые и, как правило, продаваемые программные системы на CD или DVD дисках – фильмы, учебники, энциклопедии, игры, книги, виртуальные музеи, путеводители, рекламные материалы и т. д.

Остановимся на некоторых аппаратных средствах поподробнее.



Звуковые Карты

С течением времени перечень задач выполняемых на ПК вышел за рамки просто использования электронных таблиц или текстовых редакторов. Компакт-диски со звуковыми файлами, подготовка мультимедиа презентаций, проведение видеоконференций и телефонные средства, а также игры и прослушивание аудио CD для всего этого необходимо чтобы звук стал неотъемлемой частью ПК. Для этого необходима звуковая карта. Наличие дисководов CD-ROM позволяет только прослушивать звуковые компакт-диски. Звуковая карта необходима, чтобы получить профессиональное качество звукового сопровождения, создавать и записывать звуки, синтезировать сложные аудиоэффекты, смешивать звуковую информацию от нескольких источников, самостоятельно включать звуковое сопровождение в мультимедийные презентации, дополнять документы голосовыми аннотациями и др. Звуковая карта (или Sound Blaster) устанавливается, как правило, в виде электронной платы в разъем материнской платы компьютера. История развития звуковых карт начинается с выпуском самых первых моделей компьютеров фирмы IBM. Изначально компьютеры были снабжены только PC Speaker-ом (Динамиком), который не предназначен для воспроизведения нормального звука. В то время ни о каких мультимедиа программах никто и не помышлял, и практически единственным применением хорошему звуку были компьютерные игры. Первой звуковую карту для IBM PC сделала фирма TANDY. Звучание этой карты было примерно аналогично звучанию карт в игровых компьютерах (3 музыкальных голоса). Звуковые карты стали

выпускать и другие фирмы. Наиболее известные из них Adlib, Creative и Roland. Звуковые карты Adlib - это одни из самых простых и дешёвых из использующихся на сегодняшний день звуковых карт. Их звучание основано, как и у большинства распространённых карт, на модуляции частоты (Frequency Modulation — FM). Карта содержит 11 голосов FM и может достаточно неплохо воспроизводить музыку. Но цифровой (или оцифрованный) звук ей не под силу. Карты этого типа - монофонические, но они часто используют принцип псевдостерео. На сегодняшний день карты, поддерживающие только стандарт Adlib, морально устарели. Звуковые карты фирмы Creative Labs стали стандартом для современных систем мультимедиа. Важным отличием этих карт является возможность воспроизводить цифровой звук, то есть практически все, что только можно услышать. Качество же воспроизведения зависит от двух важнейших параметров - разрядности и частоты дискретизации (это частота обновления данных). Разрядность карт обычно составляет 8 и 16 бит, а частота дискретизации — 4 до 44.1 кГц. Цифровой звук карты воспроизводят и записывают благодаря цифро-аналоговому и аналого-цифровому преобразователям (ЦАП и АЦП). Файлы, содержащие видеоизображения и звук, имеют расширения *.avi, *.mov, *.mpg. Специальный фонограф — Sound Recorder, предназначен для записи и воспроизведения звука, а также для редактирования звуковых файлов. Звуковые файлы имеют расширения *.wav, *.mid, *.mod, *.voc, *.fli.

Лазерные диски, CD-ROM



В связи с ростом объемов и сложности программного обеспечения, широким внедрением мультимедиа приложений, сочетающих движущиеся изображения, текст и звук, огромную популярность в последнее время приобрели устройства для чтения компакт- дисков CD-ROM. Компакт- диски изначально разработанные для любителей высоко качественного звучания, прочно вошли на рынок компьютерных устройств. Оптические компакт- диски перешли на смену виниловым в 1982 году. Было решено, что стандарт рассчитан на 74 минуты звучания "Red Book". Когда 74 минуты пересчитали в байты получилось 640 Мбайт. Первые приводы имели единичную скорость (Single speed) равную 150 Кбайт/с. Модели накопителей с удвоенной скоростью появились в 1992 году. Приводы с утроенной и с учетверенной скоростью в начале 1994 году. Сегодня речь

уже идет о скорости увеличенной в шесть и даже восемь раз. Коэффициент увеличения скорости не обязательно целый. Как и в компакт-дисках, применяемых в бытовых CD-плеерах, информация на компьютерных компакт-дисках кодируется посредством чередования отражающих и не отражающих свет участков на подложке диска. При промышленном производстве компакт-дисков эта подложка выполняется из алюминия, а не отражающие свет участки делаются с помощью продавливания углублений в подложке специальной пресс-формой. При единичном производстве компакт-дисков (так называемых CD-R дисков, см. ниже) подложка выполняется из золота, а нанесение информации на нее осуществляется лучом лазера. В любом случае сверху от подложки на компакт-диске находится прозрачное покрытие, защищающее занесенную на компакт-диск информацию от повреждений. Хотя по внешнему виду и размеру используемые в компьютерах компакт-диски не отличаются от дисков, применяемых в бытовых CD плеерах, однако компьютерные устройства для чтения компакт-дисков стоят существенно дороже. Это не удивительно, ведь чтение программ и компьютерных данных должно выполняться с гораздо высокой надежностью, чем та, которая достаточна при воспроизведении музыки. Поэтому чтение используемых в компьютере компакт-дисков осуществляется с помощью луча лазера небольшой мощности. Однако скорость чтения данных с компакт-дисков значительно ниже, чем с жестких дисков. Одна из причин этого состоит в том, что компакт-диски при чтении вращаются не с постоянной угловой скоростью, а так, чтобы обеспечить неизменную линейную скорость отхождения информации под читающей головкой. Стандартная скорость чтения данных с компакт-дисков всего 150-200 Кбайт/с, а время доступа 0,4 с. Впрочем, в последнее время выпускаются в основном устройства с двойной, тройной и даже четвертой скоростью вращения, они обеспечивают соответственно более высокие скоростные показатели: время доступа 0,2-0,3 с, скорость считывания 500 Кбайт/с. Заметим, однако, что устройства с тройной скоростью в реальных задачах увеличивают скорость работы с компакт-диск не в полтора и не в два раза по сравнению с устройством с двойной скоростью, а всего на 30 - 60%. Дисководы для CD производят такие известные фирмы, как Sony, NEC, Panasonic, Plextor, Creative, LG и др.

Видеокарты



При смешении сигналов основные проблемы возникают с видеоизображением. Различные ТВ-стандарты, существующие в мире (NTSC, PAL, SECAM), применение разных мониторов и видеоконтроллеров диктует разнообразие подходов в разрешении возникающих проблем. Однако в любом случае требуется синхронизация двух изображений, для чего служит устройство генлок (genlock). С его помощью на экране монитора могут быть совмещены изображение, сгенерированное компьютером (анимированная или неподвижная графика, текст, титры), и “живое” видео. Если добавить еще одно устройство — кодер (encoder), компьютерное изображение может быть преобразовано в форму ТВ-сигнала и записано на видеопленку. “Настольные видеостудии”, являющиеся одним из примеров применения систем мультимедиа, позволяют готовить совмещенные видео-компьютерные клипы, титры для видеофильмов, помогают при монтаже кинофильмов.

Системы такого рода не позволяют как-то обрабатывать или редактировать само аналоговое изображение. Для того, чтобы это стало возможным, его необходимо оцифровать и ввести в память компьютера. Для этого служат так называемые платы захвата (capture board, frame grabbers). Оцифровка аналоговых сигналов порождает огромные массивы данных. Так, кадр стандарта NTSC (525 строк), преобразованный платой типа Truevision, превращается в компьютерное изображение с разрешением 512x482 пиксель. Если каждая точка представлена 8 битами, то для хранения всей картинке требуется около 250 Кбайт памяти, причем падает качество изображения, так как обеспечивается только 256 различных цветов. Считается, что для адекватной передачи исходного изображения требуется 16 млн. оттенков, поэтому используется 24-битовый формат хранения цветной картинке, а необходимый размер памяти возрастает. Оцифрованный кадр может затем быть изменен, отредактирован обычным графическим редактором, могут быть убраны или добавлены детали, изменены цвета, масштабы, добавлены спецэффекты, типа мозаики, инверсии и т.д. Естественно, интерактивная экранная обработка возможна лишь в пределах разрешения, обеспечиваемого данным конкретным видеоадаптером. Обработанные кадры могут быть записаны на диск в каком-либо графическом формате и затем использоваться в качестве реалистического

неподвижного фона для компьютерной анимации. Возможна также покадровая обработка исходного изображения и вывод обратно на видеопленку для создания псевдореалистического мультфильма. Запись последовательности кадров в цифровом виде требует от компьютера больших объемов внешней памяти: частота кадров в американском ТВ-стандарте NTSC — 30 кадров/с (PAL, SECAM — 25 кадров/с), так что для запоминания одной секунды полноцветного полноэкранного видео требуется 20–30 Мбайт, а оптический диск емкостью 600 Мбайт вместит менее полминуты изображения. Но последовательность кадров недостаточно только запомнить, ее надо еще вывести на экран в соответствующем темпе. Подобной скоростью передачи информации — около 30 Мбайт / с — не обладает ни одно из существующих внешних запоминающих устройств. Чтобы выводить на экран компьютера оцифрованное видео, приходится идти на уменьшение объема передаваемых данных, (вывод уменьшенного изображения в небольшом окне, снижение частоты кадровой развертки до 10–15 кадров / с, уменьшение числа бит / пиксель), что, в свою очередь приводит к ухудшению качества изображения. Более радикально обе проблемы — памяти и пропускной способности — решаются с помощью методов сжатия / развертки данных, которые позволяют сжимать информацию перед записью на внешнее устройство, а затем считывать и разворачивать в реальном режиме времени при выводе на экран. Так, для движущихся видеоизображений существующие адаптивные разностные алгоритмы могут сжимать данные с коэффициентом порядка 100:1— 160:1, что позволяет разместить на CD-ROM около часа полноценного озвученного видео. Работа этих алгоритмов основана на том, что обычно последующий кадр отличается от предыдущего лишь некоторыми деталями, поэтому, взяв какой-то кадр за базовый, для следующих можно хранить только относительные изменения. При значительных изменениях кадра, например, при монтажной склейке, наезде или панорамировании камеры, автоматически выбирается новый базовый кадр. Для статических изображений коэффициент сжатия, естественно, ниже — порядка 20–30:1. Для аудиоданных применяют свои методы компрессии.

При использовании специальных видео-адаптеров (видеобластеров) мультимедиа-ПК становятся центром бытовой видео-системы, конкурирующей с самым совершенным телевизором. Новейшие видеоадаптеры имеют средства связи с источниками телевизионных сигналов и встроенные системы захвата кадра (компрессии / декомпрессии видеосигналов) в реальном масштабе времени, т.е. практически мгновенно. Видеоадаптеры имеют быструю видеопамять до 512 Мбайт и специальные графические 3D-ускорители процессоры. Это позволяет получать до 100 кадров в секунду и обеспечить вывод подвижных

полноэкранных изображений. Имеется большое количество устройств, предназначенных для работ с видеосигналами на IBM PC совместимых компьютерах. Условно можно разбить на несколько групп: устройства для ввода и захвата видеопоследовательностей (Cupture play), фреймграбберы (Framegrabber), TV-тюнеры, преобразователи сигналов VGA-TV и др..

TVтюнеры



Эти устройства выполняются обычно в виде карт или бокса (небольшой коробочки). Они преобразуют аналоговый видеосигнал поступающий по сети кабельного телевидения или от антенны, от видеомagniтофона или камкордера (camcorder). TV-тюнеры могут входить в состав других устройств таких как MPEG-плейеры или фреймграбберы. Некоторые из них имеют встроенные микросхемы для преобразования звука. Ряд тюнеров имеют возможность для вывода телетекста.

Фреймграбберы



Как правило они объединяют графические, аналогово-цифровые и микросхемы для обработки видеосигналов, которые позволяют дискретизировать видеосигнал, сохранять отдельные кадры изображения в буфере с последующей записью на диск либо выводить их непосредственно в окно на мониторе компьютера. Содержимое буфера обновляется каждые 40 мс. То есть с частотой смены кадров. Вывод видеосигналов происходит в режиме наложения (overby). Для реализации окна на экране монитора с "живым" видео карта фреймграббера соединена с графическим адаптером через 26 контактный Feature коннектор. С ним обычно поставляется пакет Video for Windows вывод картинки размером 240*160 пикселей при воспроизведении 256 цветов и больше. Первые устройства - Video Blaster, Video Spigot.

Преобразователи

VGA-TV

Данные устройства транслируют сигнал в цифровом образе VGA изображения в аналоговый сигнал пригодный для ввода на телевизионный приемник. Производители обычно предлагают подобные устройства выполненные либо как внутренние ISA карта либо как внешний блок. Ряд преобразователей позволяют накладывать видеосигнал например для создания титров. При этом осуществляется полная синхронизация преобразованного компьютерного сигнала по внешнему (gtnlok). При наложении формируется специальный ключевой (key) сигнал трех видов lumakey, chromakey или alpha chenol.

MPEG-плееры

Данные устройства позволяют воспроизводить последовательности видеоизображения (фильмы) записываемых на компакт- дисках, качеством VNS. Скорость потока сжатой информации не превышает обычно 150 Кбайт/с. Основная сложность задачи решаемой MPEG кодером, состоит в определении для каждого конкретного видеопотока оптимального соотношения между тремя видами изображения: (I)ntra, (P)redicted и (B)idirectional. Первым MPEG –плеером была плата Reel Magic компании Sigina Desing в 1993 году.

Задания для самостоятельной работы

1. Оценить информационный объем моноаудиофайла длительностью звучания 10 с. при 16 битном кодировании и частоте дискретизации 44 кГц.
2. Используются графические режимы с глубинами цвета 8, 16, 24 и 32 бита. Вычислить объемы видеопамати, необходимые для реализации данных глубин цвета при различных разрешающих способностях экрана. (800×600, 1024 × 768, 1280 × 1024)