

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Амурский государственный университет»

МУЛЬТИМЕДИА ТЕХНОЛОГИИ

сборник учебно-методических материалов для направления подготовки:
10.03.01 «Информационная безопасность»,
направленность (профиль) «Безопасность автоматизированных систем»
(по отрасли или в сфере профессиональной деятельности)

2019 г.

Печатается по решению
редакционно-издательского совета
факультета математики и информатики
Амурского государственного
университета

Составитель: Назаренко Н.В.

Мультимедиа-технологии: сборник учебно-методических материалов для направления подготовки 10.03.01 «Информационная безопасность», направленность (профиль) «Безопасность автоматизированных систем» (по отрасли или в сфере профессиональной деятельности) – Благовещенск: Амурский гос. ун-т, 2019 г.

© Амурский государственный университет, 2019

©Кафедра информационной безопасности, 2019

© Назаренко Н.В., составление

СОДЕРЖАНИЕ

1. Краткое изложение лекционного материала	4
2. Методические рекомендации (указания) к лабораторным занятиям	66
2. Методические указания для самостоятельной работы студентов	77

1. КРАТКОЕ ИЗЛОЖЕНИЕ ЛЕКЦИОННОГО МАТЕРИАЛА

РАЗДЕЛ 1. ОБЩИЕ СВЕДЕНИЯ О МУЛЬТИМЕДИЙНЫХ ТЕХНОЛОГИЯХ

Тема 1.1. Классификация и область применения мультимедийных приложений

1.1. Составляющие части мультимедиа

Слово «мультимедиа» прочно вошло в наш лексикон, и без него уже трудно представить современный компьютерный мир. Как и всякое удачное понятие, оно многообразно. Пожалуй, наиболее точная формулировка принадлежит одному из пионеров мультимедиа в нашей стране Сергею Новосельцеву: «Мультимедиа (англ. multimedia от лат. multum – много и media, medium – средоточие, средства) – это комплекс аппаратных и программных средств, позволяющих пользователю работать в диалоговом режиме с разнородными данными (графикой, текстом, звуком, видео и анимацией), организованными в виде единой информационной среды». Как видим, мультимедиа объединяет несколько типов разнородных данных (текст, звук, видео, графическое изображение и анимацию) в единое целое. И это понятие само по себе имеет три лица.

Во-первых, мультимедиа – как идея, т. е. новый подход к хранению информации различного типа в единой цифровой форме. Во-вторых, мультимедиа – как оборудование для обработки и хранения информации, без него мультимедиа-идею реализовать невозможно. В-третьих, это программное обеспечение, позволяющее объединить четыре элемента информации в законченное мультимедиа-приложение.

Мультимедиа технологии являются одним из наиболее перспективных и популярных направлений информатики. Они имеют целью создание продукта, содержащего «коллекции изображений, текстов и данных, сопровождающихся звуком, видео, анимацией и другими визуальными эффектами (Simulation), включающего интерактивный интерфейс и другие механизмы управления». Данное определение сформулировано в 1988 году крупнейшей Европейской Комиссией, занимающейся проблемами внедрения и использования новых технологий. Интерактивность – свойство реагировать на действия пользователей, в том числе и управлять пользователем. Идейной предпосылкой возникновения технологии мультимедиа считают концепцию организации памяти «МЕМЕХ», предложенную еще в 1945 году американским ученым Ваннивером Бушем. Она предусматривала поиск информации в соответствии с ее смысловым содержанием, а не по формальным признакам (по порядку номеров, индексов или по алфавиту и т. п.). Эта идея нашла свое выражение и компьютерную реализацию сначала в виде системы гипертекста (система работы с комбинациями текстовых материалов), а затем и гипермедиа (система, работающая с комбинацией графики, звука, видео и анимации), и, наконец, в мультимедиа, соединившей в себе обе эти системы.

Однако всплеск интереса в конце 80-х годов XX в. к применению мультимедиа технологии в гуманитарной области (и, в частности, в историко-культурной) связан, несомненно, с именем выдающегося американского компьютерщика-бизнесмена Билла Гейтса, которому принадлежит идея создания и успешной реализации на практике мультимедийного продукта на основе служебной музейной инвентарной базы данных с использованием в нем всех возможных «сред»: изображений, звука, анимации, гипертекстовой системы («National Art Gallery London»). Именно этот продукт аккумулировал в себе три основных принципа мультимедиа:

- представление информации с помощью комбинации множества воспринимаемых человеком сред;
- наличие нескольких сюжетных линий в содержании продукта (в том числе и выстраиваемых самим пользователем на основе «свободного поиска» в рамках предложенной в содержании продукта информации);
- художественный дизайн интерфейса и средств навигации. Несомненным достоинством и особенностью технологии являются следующие возможности мультимедиа, которые активно используются в представлении информации:

- возможность хранения большого объема самой разной информации на одном носителе;
- возможность увеличения (детализации) на экране изображения или его наиболее интересных фрагментов, иногда в двадцатикратном увеличении (режим «лупа») при сохранении качества изображения. Это особенно важно для презентации произведений искусства и уникальных исторических документов;
- возможность сравнения изображения и обработки его разнообразными программными средствами с научно-исследовательскими или познавательными целями;
- возможность выделения в сопровождающем текстовом или другом визуальном материале «горячих слов (областей)», по которым осуществляется немедленное получение справочной или любой другой пояснительной (в том числе визуальной) информации (технологии гипертекста и гипермедиа);
- возможность осуществления непрерывного музыкального или любого другого аудиосопровождения, соответствующего статичному или динамичному визуальному ряду;
- возможность использования видеофрагментов из фильмов, видеозаписей и т.д., функции «стоп-кадра», покадрового «пролистывания» видеозаписи;
- возможность включения в содержание диска баз данных, методик обработки образов, анимации (к примеру, сопровождение рассказа о композиции картины графической анимационной демонстрацией геометрических построений ее композиции) и т.д.; возможность подключения к глобальной сети Internet;
- возможность работы с различными приложениями (текстовыми, графическими и звуковыми редакторами, картографической информацией);
- возможность создания собственных «галерей» (выборки) из представляемой в продукте информации (режим «карман» или «мои пометки»);
- возможность «запоминания пройденного пути» и создания «закладок» на заинтересовавшей экранной «странице»;
- возможность автоматического просмотра всего содержания продукта («слайд-шоу») или создания анимированного и озвученного «путеводителя гида» по продукту («говорящей и показывающей инструкции пользователя»), включение в состав продукта игровых компонентов с информационными составляющими;
- возможность «свободной» навигации по информации и выхода в основное меню (укрупненное содержание), на полное оглавление или вовсе из программы в любой точке продукта.

Итак, мультимедийный продукт – наиболее эффективная форма подачи информации в среде компьютерных информационных технологий. Он позволяет собрать воедино огромные и разрозненные объемы информации, дает возможность с помощью интерактивного взаимодействия выбирать интересующие в данный момент информационные блоки, значительно повышая эффективность восприятия информации.

Мультимедиа (от англ. multi - много, media - среда) – комбинированное представление информации в разных формах (текст, звук, видео и т. д.).

Технология мультимедиа - интерактивная технология, обеспечивающая работу с неподвижными изображениями, видеоизображением, анимацией, текстом и звуковым рядом.

Развитием гипертекстовых технологий в глобальных сетях стало появление гипермедийных документов, которые наряду с текстовой информацией содержат информацию, представленную в мультимедийной форме.

1.2. Классификация мультимедиа приложений

Мультимедиа – это взаимодействие визуальных и аудиоэффектов под управлением интерактивного программного обеспечения. Мультимедиа – комбинация текста, графических изображений, звука, анимации и видеоэлементов. Согласно представленным выше определениям, мультимедиа можно классифицировать с разных точек зрения: – на основе поддержки взаимодействия, – на основе использования различных мультимедийных телекоммуникационных технологий.

Мультимедиа информация содержит не только традиционные статистические элементы: текст, графику, но и динамические: видео-, аудио- и анимационные последовательности. Типы данных мультимедиа информации представлены на рис. 1.1.



Рисунок 1.1 Типы данных мультимедиа информации

1.3. Области применения мультимедиа приложений

Области применения:

- Обучение с использованием компьютерных технологий (научнопросветительская или образовательная сфера);
- Видеоэнциклопедии, интерактивные путеводители, тренажеры, ситуационно-ролевые игры и др.;
- Информационная и рекламная служба;
- Популяризаторская и развлекательная сферы;
- Интернет-вещание;
- Развлечения, игры, системы виртуальной реальности;
- Презентационная (витринной рекламы), СМИ;
- Творчество (станция мультимедиа становится незаменимым авторским инструментом в кино и видеоискусстве. Автор фильма за экраном такой настольной системы собирает, «аранжирует», создает произведения из заранее подготовленных – нарисованных, отснятых, записанных и т. п. – фрагментов;
 - Военные технологии;
 - Промышленность и техника (сенсорные экраны);
 - Торговля.
- В научно-исследовательской области – это электронные архивы и библиотеки – для документирования коллекций источников и экспонатов, их каталогизации и научного описания, для создания «страховых копий», автоматизации поиска и хранения, для хранения данных о местонахождении источников для хранения справочной информации, для обеспечения доступа к внемузейным базам данных, для организации работы ученых не с самими документами, а с их электронными копиями и т.д.).
 - Медицина: базы знаний, методики операций, каталоги лекарств и т.п.
 - Искусственный интеллект – внедрение элементов искусственного интеллекта в системе мультимедиа. Они обладают способностью «чувствовать» среду общения, адаптироваться к ней и оптимизировать процесс общения с пользователем: они подстраиваются под читателей, анализируют круг их интересов, помнят вопросы, вызывающие затруднения и могут сами предложить дополнительную или разъясняющую информацию.
 - Системы распознавания речи, понимающие естественный язык, еще более расширяют диапазон взаимодействия с компьютером.

1.4. Аппаратные средства мультимедиа технологии

Технологию мультимедиа составляют специальные аппаратные и программные средства. 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 8 дисплея, декодеры для взаимного преобразования телевизионных стандартов, специальные интегральные схемы для сжатия данных в файлы допустимых размеров и т.д.

Технологию *мультимедиа* составляют две основные компоненты - аппаратная и программная, представленные на рис. 1.2.

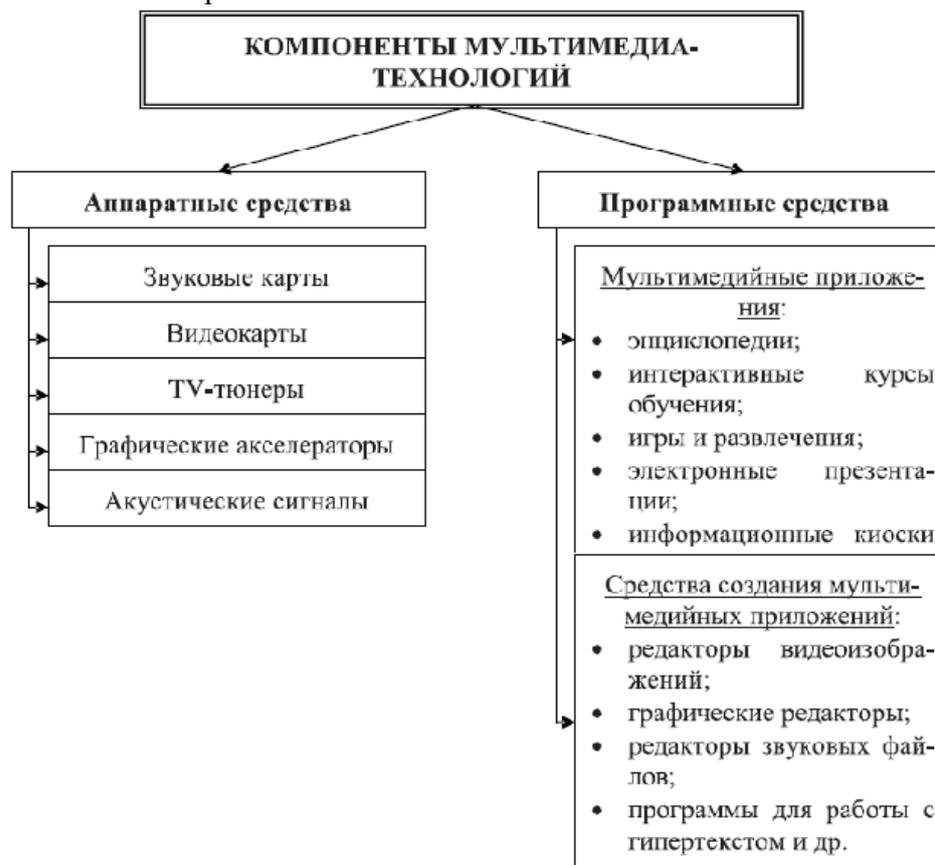


Рисунок 1.2 Компоненты мультимедиа технологий

1.4.1. 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 гарнитур);
- 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.

1.5. Программные средства мультимедиа технологий

Программные средства мультимедиа складываются из трех компонентов:

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

1.5.1. Системные 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 харак-

теристиками 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.

1.5.2. Инструментальные программные средства

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.

1.5.3. Прикладные программные средства

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.

Формат GIF (Graphics Interchange Format - формат обмена изображениями) - один из старейших форматов записи изображений. Он был разработан в 1978 г. Формат GIF рассчитан на табличное кодирование изображений с применением 256-цветной палитры, при котором одним байтом записывают одно значение некоторого произвольного цвета. Для уменьшения объема полученные данные в процессе записи сжимаются по определенным алгоритмам. Этот формат используют для представления малоцветных изображений, имеющих большие области одинакового цвета.

Формат JPEG (Joint Photographic Experts Group - объединенная экспертная группа по записи изображений) является международным стандартом. Этот формат предназначен для эффективной записи полноцветных графических изображений. Он использует наличие необязательных данных в графических изображениях, например, для случайного просмотра человеческим глазом не требуется высокого разрешения для цветовой информации в изображении. Поэтому данные, представляющие высокое цветовое разрешение, могут быть исключены. Особенностью формата JPEG является использование схемы "кодирование с потерями", т. е. при воспроизведении данных, записанных со сжатием в формате JPEG, полученная последовательность неточно соответствует данным, имевшимся перед записью.

Запись и кодирование видеоизображений основано на представлении видеоряда в виде последовательности кадров и кодировании каждого из них как отдельного изображения с последующей записью последовательности кадров. Одним из наиболее распространенных методов кодирования видеоизображений является метод MPEG (Moving Picture Experts Group - Экспертная группа по записи видеоизображений).

Базовым объектом кодирования в стандарте MPEG является кадр телевизионного изображения. Поскольку большинство кадров имеют сравнительно небольшие отличия друг от друга, то, записав полностью один кадр, можно при записи последующего кадра записывать не все изображение, а только его отличия от предыдущего. В общей последовательности видеоряда выделяют опорные и промежуточные кадры. Опорные кадры являются начальными кадрами новых сцен, а промежуточные соответствуют одной сцене и имеют много общего с опорными кадрами.

Кодирование видеоряда начинается с записи опорного кадра. Сначала изображение

разбивается на блоки фиксированного размера, которые кодируются и сжимаются с использованием специальных алгоритмов. Следующий кадр тоже разбивается на аналогичные блоки, которые сравниваются с блоками опорного кадра, а затем записываются только отличия от предыдущего кадра.

Существуют несколько разновидностей формата записи MPEG: MPEG-1, MPEG-2, MPEG-4, которые отличаются друг от друга качеством записи и степенью сжатия.

Звуковые сигналы характеризуются двумя величинами: частотой (высота звука) и амплитудой (громкость звука). Основным стандартным форматом записи звука является формат WAV, введенный в действие компаниями IBM и Microsoft. Существуют и другие форматы звуковых файлов, введенные другими корпорациями, однако выборки данных при звукозаписи имеют огромные размеры. Для передачи звука и музыки по медленным каналам связи, таким как телефонные соединения, используемые для доступа к Internet, используют специальный формат записи MP3 (MPEG-1 layer 3). В его основу положены особенности человеческого слухового восприятия, выражающиеся в том, что далеко не вся информация, которая содержится в звуковом сигнале, является полезной и необходимой - большинство слушателей ее не воспринимают. Поэтому определенная часть данных может быть сочтена избыточной. Эта "лишняя" информация удаляется без особого вреда для субъективного восприятия. При этом стандарт позволяет в заданных пределах менять параметры кодирования - получать меньшую степень сжатия при лучшем качестве или, наоборот, идти на потери в восприятии ради более высокого коэффициента компрессии.

Аппаратные средства мультимедиа включают 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. Все оборудование, отвечающее за преобразование звуковых сигналов, объединяют в звуковые карты, а за преобразование видеосигналов - в видеокарты.

Звуковая карта - это плата, микросхема, позволяющая записывать и воспроизводить звуки, синтезировать музыку, управлять внешней акустической аппаратурой, подключенной к компьютеру.

Видеокарта - это плата, микросхема, согласующая обмен графической информацией между центральным процессором и дисплеем и управляющая выводом информации на экран.

Кроме того, существуют узкоспециальные аппаратные средства, выполняющие отдельные функции: TV-тюнеры для преобразования телевизионных сигналов; графические акселераторы (ускорители), в том числе для поддержки трёхмерной графики; акустические системы с наушниками или динамиками и др.

Программные средства мультимедиа включают:

мультимедийные приложения - энциклопедии, интерактивные курсы обучения по всевозможным предметам, игры и развлечения, работа с Internet, тренажеры, средства торговой рекламы, электронные презентации, информационные киоски, установленные в общественных местах и предоставляющие различную информацию и др.

средства создания мультимедийных приложений - редакторы видеоизображений; профессиональные графические редакторы; средства для записи, создания и редактирования звуковой информации, позволяющие подготавливать звуковые файлы для включения в программы, изменять амплитуду сигнала, накладывать или убирать фон, вырезать или вставлять блоки данных на каком-то временном отрезке; программы для манипуляции с сегментами изображений, изменения цвета, палитры; программы для реализации гипертекстов и др.

Тема 1.2. Каналы передачи видеoinформации. Стандарты хранения видеоданных

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. Когда в быту говорят «видео» — то обычно имеют в виду видеоматериал, телесигнал или кинофильм, записанный на физическом носителе (видеокассете, видеодиске и т. п.).

Количество (частота) кадров в секунду — это число неподвижных изображений, сменяющих друг друга при показе 1 секунды видеоматериала и создающих эффект движения объектов на экране. Чем больше частота кадров в секунду, 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). В традиционном плёночном кинематографе используется частота 24 кадра в секунду. Системы телевидения PAL и SECAM используют 25 кадров в секунду (англ. 25 fps или 25 Герц), а система NTSC использует 29,97 кадров в секунду. Компьютерные оцифрованные видеоматериалы хорошего качества, как правило, используют частоту 30 кадров в секунду.

Оборудование для обработки видео на компьютере.

Для записи видеoinформации необходимо:

специальная плата или устройство для оцифровки видеоизображения;

видеомагнитофон или видеокамера;

программное обеспечение для записи и редактирования цифрового видео.

звуковая карта (если плата видеозахвата не поддерживает возможности захвата звука).

Видеокарта (видеоадаптер). Совместно с монитором видеокарта образует видеоподсистему персонального компьютера. Физически видеоадаптер выполнен в виде отдельной дочерней платы, которая вставляется в один из слотов материнской платы и называется видеокартой. Видеоадаптер взял на себя функции видеоконтроллера, видеопроцессора и видеопамати.

За время существования персональных компьютеров сменилось несколько стандартов видеоадаптеров: MDA (монохромный); CGA (4 цвета); EGA (16 цветов); VGA (256 цветов). В настоящее время применяются видеоадаптеры SVGA, обеспечивающие по выбору воспроизведение до 16,7 миллионов цветов с возможностью произвольного выбора разрешения экрана из стандартного ряда значений.

Плата оцифровки видео

Можно воспользоваться простейшей аналоговой картой видеозахвата или TV-тюнером. При этом существуют следующие особенности такой платы. Она должна:

показывать и захватывать аналоговое видео со скоростью потока данных, ограничиваемым только устройством записи;

захватывать видео с произвольными размерами кадра, в частности, с разрешением 352×288 (необходимое для стандарта MPEG-1);

захватывать видео как через композитный вход, так и через S- Video.

Основные форматы видео файлов

Audio Video Interleaved (*.AVI) - формат, разработанный Microsoft для записи и воспроизведения видео в операционной системе Windows. При записи в этом формате используются несколько различных алгоритмов сжатия (компрессии) видеоизображения. Среди них Cinepak, Indeo video, Motion-JPEG (M-JPEG) и др. Но только M-JPEG был признан среди них как международный стандарт для сжатия видео. Первоначально для захвата и воспроизведения видео использовались возможности программного комплекта Video for Windows, разработанного Microsoft. Компания Microsoft разработала два формата, призванных заменить формат AVI: Advanced Streaming Format (*.ASF) и Advanced Authoring Format (*.AAF).

Windows Media Video (*.WMV) - новый формат видео от Microsoft, который приходит на смену формату AVI. В его основе Windows Video Codec, разработанный на базе стандарта

MPEG-4.

Quick Time Move (*.MOV) - наиболее распространенный формат для записи и воспроизведения видео, разработанный фирмой Apple для компьютеров Macintosh в рамках технологии Quick Time. Включает поддержку не только видео, но и звука, текста, потоков MPEG, расширенного набора команд MIDI, векторной графики, панорам и объектов (QT) и трехмерных моделей. Поддерживает несколько различных форматов сжатия видео, в том числе MPEG, а также свой собственный метод компрессии.

MPEG (*.MPG, *.MPEG) - формат для записи и воспроизведения видео, разработанный группой экспертов по движущимся изображениям (MPEG). Имеет собственный алгоритм компрессии. В настоящее время активно используются для записи цифрового видео. Наиболее широкое распространение нашли два формата: MPEG-1 и MPEG-2. Они различаются по объему и качеству получаемой видеoinформации и признаны международными стандартами для сжатия видео. В настоящее время наряду с MPEG-1 и MPEG-2 используется новый формат MPEG-4. Он позволяет сжать информацию с большим коэффициентом сжатия.

Digital Video (*.DV) - формат, разработанный для цифровых видеокамер и видеоманитофонов. Кодер-декодер (кодек) определен ведущими мировыми производителями электроники, чтобы его могли поддерживать производители в своих платах с интерфейсом FireWare и комплексных решениях для редактирования цифрового видео. Формат не является компактным, поэтому необходимо его преобразование в MPEG.

Кодеки для видеoinформации

Кодек является сокращением от слов компрессор и декомпрессор, это любая технология для сжатия и обратного восстановления данных. Кодеки могут быть реализованы на уровне программного обеспечения, аппаратной части или в их комбинации. Кодеки используются для того, чтобы сделать большие видео файлы намного меньше, делая их пригодными для распространения по сети, локальной или всемирной, или любого другого метода передачи файлов.

Часть кодеков пользователь получает при установке операционной системы Windows. Это те кодеки, которые используются в файлах формата AVI. Если требуется пользоваться для записи форматом QuickTime, то нужно установить его поддержку.

Cineraк - создавался фирмой SuperMac для компьютеров с процессорами Motorola 68030 и Intel386 и с односкоростными дисковыми CD-ROM. Файлы AVI, в которых использован этот кодек, могут быть переведены в формат QuickTime и, наоборот, без переупаковки.

Sorenson Video - использует усовершенствованные алгоритмы векторного квантования и компенсации движения и адаптивное управление потоком. Оптимизирован для работы со скоростями от 2 до 100 Кбайт/с. Качество изображения значительно превышает Cineraк даже при меньших размерах файла. Сжатие очень медленное. Включен в стандартную поставку QuickTime.

Indeo Video Interactive (IVI) - видеокодек для форматов AVI и MOV требует компьютеры класса Pentium. Кодек основан на wavelet - компрессии.

MPEG - кодек, который, является официальным стандартом для сжатия и видеoinформации. В нем применяется дискретное косинусное преобразование (DCT, или ДКП) с межкадровым предсказанием. Существует несколько версий этого стандарта (MPEG-1, MPEG-2; MPEG-4).

MPEG-4 задает правила организации объектно-ориентированной среды. Он имеет дело не просто с потоками и массивами медиаданных, а с медиа объектами, например, аудио, видео, аудио/видео, графическими (плоскими и трехмерными), текстовыми. При этом MPEG-4 обеспечивает наилучшую эффективность сжатия. Используется алгоритм "сжатия повышенной эффективности" (ACE - Advanced Coding Efficiency). Можно обрабатывать звуковые сигналы в диапазоне от 2 до 24 Кбит/с, а видеосигнал - от 5 до 10 Мбит/с. Благодаря такой масштабируемости, аудио и видеоданные можно адаптировать к реальному примене-

нию. Этим обусловлен более универсальный характер MPEG-4 по сравнению с MPEG-2. Наиболее интересны практические достижения в создании видеокодеков, построенных на основе стандарта MPEG-4. Среди них - Windows Media Codec и DivX.

Программы для видеомонтажа

Чтобы превратить оцифрованную информацию в готовый продукт, ее необходимо обработать: разместить монтажные эпизоды, задать эффекты и переходы между ними, добавить титры и пояснительные тексты, отредактировать звуковое сопровождение, наконец, смонтировать готовый фильм. Для этого можно использовать следующие программы: Windows Movie Maker, Pinnacle Studio, VideoStudio, Video Wave, Media Studio Pro, Adobe Premiere, Speed Razor Pro, Adobe After Effects, Cannopus Edius и др.

Основные операции с видео файлами

Импорт

Добавление/удаление кадров

Изменение размера кадра

Разбиение кадра на фрагменты

Добавление эффектов и переходов

Тоновая коррекция кадров

Добавление интерактивных элементов управления

Добавление текста (титры, бегущая строка, текст на кадре и т.д.)

Передача видеoinформации становится все более актуальной для различных сфер жизнедеятельности - телевидения, научных исследований, медицины, современных технологий дистанционного обучения и др. В связи с активным развитием информационных технологий появились широкие технические возможности для реализации различных методов обработки и высококачественной передачи видеoinформации. Передача видеoinформации через сеть Интернет является важнейшей составляющей информационного потока для многих современных мультимедиа-приложений. При этом передача видеoinформации в реальном времени предъявляет повышенные требования к ширине полосы частот, задержкам передачи и допустимым потерям данных. В настоящее время сеть Интернет не всегда обеспечивает гарантированное качество обслуживания (Quality of Service - QoS). Кроме этого, неоднородность структуры сетей и характеристик систем передачи и приема видеoinформации затрудняет передачу в режиме распределенного доступа. Освоение наиболее подходящих стандартов сжатия, преобразования и представления видеoinформации, разработка протоколов и методов передачи составляют важную проблему в области развития информационных технологий. Для решения этой проблемы необходимо контролировать ситуации превышения пропускной способности каналов, осуществлять адаптивную подстройку характеристик сжатия и кодирования видеoinформации и обеспечивать динамическое управление уровнем допустимых ошибок передачи видеoinформации за счет предотвращения ожидаемых ошибок, дублирующей передачи данных, обеспечения малой заметности ошибок.

Использование новых возможностей информационных технологий становится более эффективным при оптимизации структуры информационных сетей, отдельных узлов и их взаимодействия. В наиболее совершенных системах критические характеристики распределяются в известной мере согласованно в структурной части, характеристиках оборудования и программном обеспечении, что повышает качество и надежность работы и расширяет функциональные возможности системы в целом.

В настоящей статье рассмотрены подходы к оптимизации условий передачи видеoinформации, предотвращения и коррекции ошибок передачи, а также выполнен анализ характеристик основных стандартов и протоколов, используемых при передаче видеoinформации в компьютерных сетях.

Особенности передачи видеoinформации в компьютерных сетях

Для обеспечения требуемого качества передачи видеoinформации в реальном времени требуется определенная ширина полосы или пропускная способность канала. Существуют

два принципа потоковой передачи звука и изображения в Интернет -одноадресная передача данных (Unicast) и многоадресная передача (Multicast).

В режиме одноадресной передачи данных (рис. 1, а) сервер-отправитель информации формирует для каждого клиента отдельный поток данных, а компьютер пользователя периодически отправляет на сервер подтверждение о доставке

информационных пакетов. При этом требуемые мощность сервера и полоса пропускания канала связи прямо пропорциональны количеству клиентов. Как видно из примера рис. 1, а, пять копий потока данных пересылаются через участок 1 и три копии - через участок 2. Одноадресная передача данных используется, в основном, в системах "видео по запросу" (video-on-demand). Она удобна для работы с видеоархивами единичных пользователей и для распространения вещания на абонентов, подключенных по низкоскоростным коммутируемым линиям связи.

В режиме многоадресной передачи данных (рис. 1, б) сервер формирует один поток данных, к которому могут подключаться по сети различные группы клиентов. Здесь мощность сервера и полоса пропускания канала не зависят от количества получателей информации. Многоадресная передача удобна, например, для Интернет-телевидения и видеоконференций, поскольку каждый пользователь имеет возможность подключения к источнику информации.

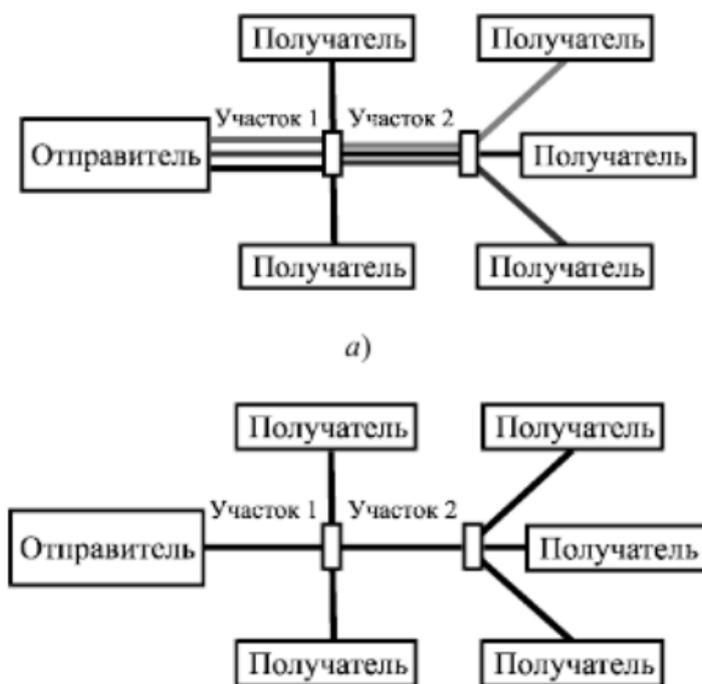


Рисунок 1. Схемы одноадресной (а) и многоадресной (б) передачи данных

Однако это преимущество достигается ценой отсутствия подтверждения о доставке информации и ручной настройки каждого маршрутизатора или коммутатора на всем сетевом пути от сервера-отправителя до группы получателей информации. Технология многоадресной передачи данных применяется для новостийных и обучающих программ в корпоративных сетях. Для российского сегмента Интернет, где ощущается дефицит высокоскоростных соединений сетевых узлов, сдерживающий широкое использование вещательных технологий, применение многоадресной передачи данных особенно актуально.

Обычные маршрутизаторы не обеспечивают контроля переполнения или превышения пропускной способности канала связи и, следовательно, гарантированного качества передачи. Другая проблема состоит в том, что при передаче непрерывной видеoinформации каждый информационный пакет должен вовремя достигать получателя, быть декодирован и отображен без пауз. Задержки передачи пакетов нарушают непрерывность отображения, что неприятно для визуального восприятия, а в ряде случаев является неприемлемым. В настоя-

щее время сеть Интернет не обеспечивает гарантированной доставки информационных пакетов в реальном времени и без потерь. Потери пакетов считаются в среднем допустимыми, если их относительный уровень составляет не более долей процента. В ситуациях превышения пропускной способности потери информации могут быть гораздо большими, и качество передачи резко снижается. Неоднородность структуры сети и характеристик отдельных систем передачи и приема приводят к изменчивости значений задержек и потерь, которые в результате не могут быть скомпенсированы. Этот недостаток особенно характерен для многоадресной передачи видеoinформации.

При решении указанных проблем могут быть использованы два основных подхода. Первый из них состоит в обеспечении требуемого QoS за счет поддержки всех необходимых параметров сети при глобальном усовершенствовании сетевого оборудования. Второй подход основан на использовании на стороне приема видеoinформации таких систем, которые обеспечивали бы приемлемое качество при малой зависимости от характеристик сети. Этот подход рассматривается как более предпочтительный, поскольку не требует кардинального изменения параметров сети.

Повышение качества передачи видеoinформации в компьютерных сетях может быть достигнуто в двух направлениях: либо за счет совершенствования методов управления, обработки и передачи безотносительно к характеру видеoinформации, либо на основе использования методов сжатия, учитывающих семантику изображений на различных уровнях компрессии.

Управление параметрами передачи видеoinформации

Управление параметрами сети осуществляется тремя основными способами: варьированием скорости передачи, адаптивным кодированием видеoinформации, характеристики которого определяются скоростью передачи, и локальным ускорением передачи на коротких интервалах за счет повышения степени сжатия. Первый способ реализуется на уровне передачи информации, второй основывается на специальных методах сжатия видеoinформации, третий включает в себя управление передачей и сжатием видеoinформации.

В качестве примера на рис. 2 представлена структурная схема управления, включающая перечисленные процедуры.

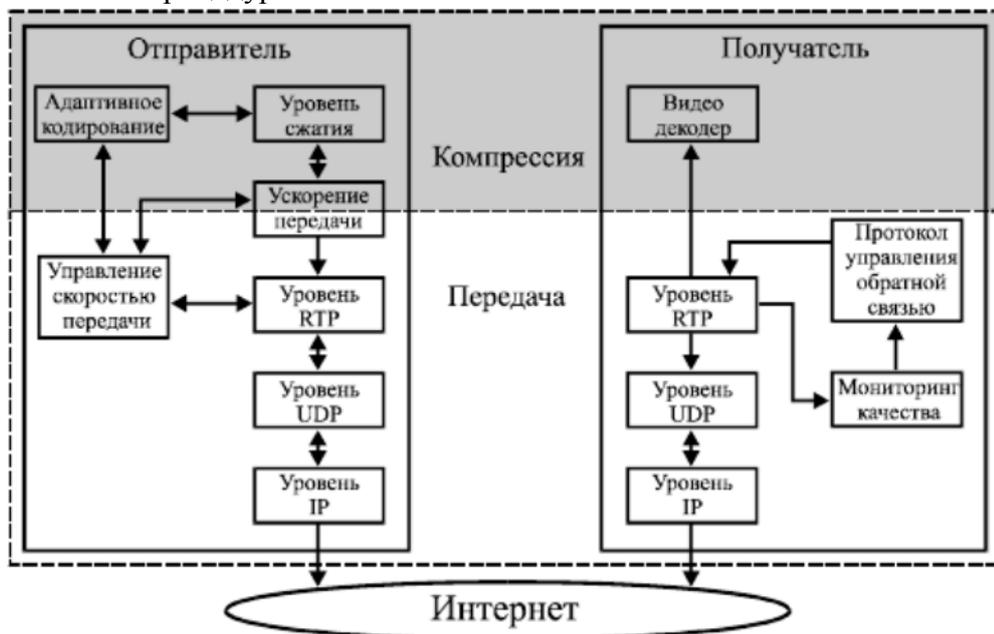


Рисунок 2. Иерархия системы передачи видеoinформации в реальном времени

Система со структурой рис. 2 ориентирована на передачу видеoinформации в реальном времени, однако она может быть использована для стационарного потока видеоданных в случае отключения блоков адаптивного кодирования в зависимости от скорости передачи. В схеме реализуется управление скоростью передачи на стороне источника информации (от-

правителя). В динамическом режиме передачи после адаптивного сжатия поток видеоданных обрабатывается в блоке локального ускорения передачи и далее преобразуется на уровнях RTP/UDP/IP перед поступлением в сеть.

Транспортный протокол RTP (Real-Time Transport Protocol) обеспечивает передачу данных между конечными узлами сети, работает поверх протокола IP и является альтернативой широко используемому протоколу TCP. Хотя протокол TCP и гарантирует доставку пакетов данных в нужной последовательности, трафик при этом очень неравномерен (пакеты испытывают всевозможные задержки). Обладая способностью распознавать содержимое пакетов (например, различать видеоданные, соответствующие спецификациям MPEG и H.261), а также средствами обнаружения потери данных, протокол RTP снижает задержки до уровня, необходимого для успешной передачи потоков интегрированных аудио-, видео- и цифровых данных. Протокол RTP обеспечивает идентификацию типа и номера пакета, устанавливает в него метку синхронизации. На основе этой информации приемный терминал синхронизирует данные и осуществляет их последовательное и непрерывное воспроизведение. Корректное функционирование RTP возможно при наличии в абонентских терминалах механизмов буферизации принимаемой информации.

Протокол UDP (User Datagram Protocol) используется в сетях с негарантированной полосой пропускания с целью минимизации задержек и максимального использования имеющейся полосы пропускания для передачи видеопотоков. Этот протокол реализует упоминавшийся выше механизм многоадресной рассылки (IP Multicast) для негарантированной доставки аудио- и видеoinформации определенному числу пользователей. Поверх IP Multicast работает RTP, который создает необходимые условия для нормального воспроизведения полученных потоков данных на абонентских терминалах. Однако протокол UDP не обеспечивает контроля превышения пропускной способности канала, поэтому такой контроль должен быть обеспечен на более высоких уровнях в схеме рис. 2.

На стороне получателя информационные пакеты преобразуются на уровнях IP/UDP/RTP перед их декодированием. В схеме рис. 2 предусмотрен динамический мониторинг показателя QoS, основывающийся на свойствах полученных пакетов, а именно на контроле потерь и задержек при их получении. Используя эту информацию, блок управления обратной связью обеспечивает оценку требуемой ширины полосы для выбора подходящего режима адаптивного кодирования и локального ускорения скорости передачи за счет локального повышения степени сжатия.

РАЗДЕЛ 2. КОМПЬЮТЕРНЫЕ АУДИОТЕХНОЛОГИИ

Тема 2.1. Звуковые системы персонального компьютера

Звук – это механические колебания среды: воздуха, воды и т.д, воспринимаемые слуховым аппаратом человека. То, что мы слышим – это результат обработки колебательных движений барабанной перепонки уха, представленный в виде сигналов нервной системы. Вне среды переноса звуковых волн звук не существует. Однако звуковые колебания можно перевести на другой носитель: изменить представление информации, не теряя ее фактически. Обычно звуковые колебания переносят на сигналы радиоволн.

Звуковую информацию можно также представить в математической форме, в виде периодических функций времени. Это представление обычно записывают в виде формулы

Здесь $A(t)$ – амплитуда звукового сигнала, а t – время.

Все слышимые звуки являются результатом воздействия звуковых волн. На магнитной ленте, виниловой пластинке звук сохраняется в виде непрерывного электрического сиг-

$$A(t) = a_0 + \sum_{k>0} (a_k \cos \omega_k t + b_k \sin \omega_k t)$$

нала, определяющего изменение звуковых волн. Звук, создаваемый электрическими волнами, называют аналоговым.

Звук может храниться на цифровых носителях, т.е. быть представленным в виде набора цифр. Любая цифровая техника или программа работают со звуком, представленным в цифровом виде. Таким образом, для переноса звука на цифровой носитель, необходимо осуществить его аналогово-цифровое преобразование. Такое преобразование состоит из трех этапов:

дискретизация – представление непрерывного сигнала в виде последовательного набора отдельных амплитуд;

квантование – разделение каждой амплитуды на заданное число уровней;

кодирование – запись данных позиции и уровня амплитуды в цифровом виде.

На практике преобразования звуковой информации из непрерывной формы в дискретную выполняются электронными устройствами, называемыми аналого-цифровыми преобразователями (АЦП) и цифро-аналоговыми преобразователями (ЦАП). Современные звуковые карты могут обеспечить кодирование 65536 различных уровней сигнала или состояний. Для определения количества бит, необходимых для кодирования, решим показательное уравнение: $65536 = 2^I$, т.к. $65536 = 2^{16}$, то $I = 16$ бит.

Таким образом, современные звуковые карты обеспечивают 16-битное кодирование звука. При каждой выборке значению амплитуды звукового сигнала присваивается 16-битный код.

Количество выборок в секунду может быть в диапазоне от 8000 до 48000, т.е. 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 моно-аудио-файла длительностью звучания 1 секунду при среднем качестве звука (16 бит, 24 КГц). Для этого количество бит на одну выборку необходимо умножить на количество выборок в 1 секунду:

$$16 \text{ бит} * 24000 = 384000 \text{ бит} = 48000 \text{ байт} \approx 47 \text{ Кбайт.}$$

Устройства для работы со звуком

Звуковая карта явилась одним из наиболее поздних усовершенствований персонального компьютера. Она подключается к одному из слотов материнской платы в виде дочерней карты и выполняет вычислительные операции, связанные с обработкой звука, речи, музыки. Звук воспроизводится через внешние звуковые колонки, подключаемые к выходу звуковой карты. Специальный разъем позволяет отправить звуковой сигнал на внешний усилитель. Имеется также разъем для подключения микрофона, что позволяет записывать речь или музыку и сохранять их на жестком диске для последующей обработки и использования.

Основным параметром звуковой карты является разрядность, определяющая количество битов, используемых при преобразовании сигналов из аналоговой в цифровую форму и наоборот. Чем выше разрядность, тем меньше погрешность, связанная с оцифровкой, тем выше качество звучания. Минимальным требованием сегодняшнего дня являются 16 разрядов, а наибольшее распространение имеют 32-разрядные и 64-разрядные устройства.

Форматы звуковых файлов.

Звук в компьютере хранится в файлах, имеющих различные способы представления информации. Перечислим основные форматы хранения звуковой информации.

WAVE (*.wav) – наиболее широко распространенный звуковой формат. Используется операционной системой Windows для хранения звуковых файлов. В его основе лежит формат RIFF (Resource Interchange File Format), позволяющий сохранять данные в структурированном виде.

Стандарт MPEG-1 представляет собой, целый комплект аудио и видео стандартов. Согласно стандартам ISO (International Standards Organization), аудио часть MPEG-1 включает в себя три алгоритма различных уровней сложности: Layer 1 (уровень 1), Layer 2 (уровень 2) и Layer 3 (уровень 3). Общая структура процесса кодирования одинакова для всех уровней MPEG-1 . Вместе с тем, несмотря на схожесть уровней в общем подходе к кодированию,

уровни различаются по целевому использованию и задействованным в кодировании внутренним механизмам. Для каждого уровня определен свой формат записи выходного потока данных и, соответственно, свой алгоритм декодирования.

MPEG Layer 3 (*.mp3) - формат звуковых файлов с потерями качества, разработанный для сохранения звуков, отличных от человеческой речи. Используется для оцифровки музыкальных записей.

Windows Media Audio (*.wma) - формат звуковых файлов, предложенный фирмой Microsoft. Кодек Windows Media Audio 8 обеспечивает качество, аналогичное mp3, при размерах файлов втрое меньших.

MIDI (*.mid) - цифровой интерфейс музыкальных инструментов (Musical Instrument Digital Interface). MIDI определяет обмен данными между музыкальными и звуковыми синтезаторами разных производителей. Интерфейс MIDI представляет собой протокол передачи музыкальных нот и мелодий. Но данные MIDI не являются цифровым звуком: это сокращенная форма записи музыки в числовой форме.

Программное обеспечение для редактирования звука.

Наиболее известными в настоящее время являются следующие программы для обработки звука: Sound Forge, GoldWave, Adobe Audition и др.

Основные операции со звуком.

Запись.

Добавление/удаление звуковой дорожки.

Изменение размера звуковой дорожки.

Разбиение звуковой дорожки на фрагменты.

Редактирование звуковой кривой.

Изменение громкости звучания.

В последнее время возможности мультимедийного оборудования претерпели значительный рост, однако этой области почему-то не уделяется достаточно внимания. рядовой пользователь страдает от нехватки информации и вынужден учиться лишь на собственном опыте и ошибках. Этой статьей мы постараемся устранить это досадное недоразумение. Данная статья ориентирована на рядового пользователя и ставит своей целью помочь ему разобраться в теоретических и практических основах цифрового звука, выявить возможности и основные приемы его использования.

Что именно мы знаем о звуковых возможностях компьютера, кроме того, что в нашем домашнем компьютере установлена звуковая плата и две колонки? К сожалению, вероятно из-за недостаточности литературы или по каким-либо другим причинам, но пользователь, чаще всего, не знаком ни с чем, кроме встроенного в Windows микшера аудио входов/выходов и Recorder'а. Единственное использование звуковой карты, которое находит простой пользователь – это вывод звука в играх, да прослушивание коллекции аудио. А, ведь, даже самая простая на сегодняшний день звуковая плата, установленная почти в каждом компьютере, умеет намного больше - она открывает широчайшие возможности для всех, кто любит и интересуется музыкой и звуком, а для тех, кто хочет создавать свою музыку, звуковая карта может стать всемогущим инструментом. Для того чтобы узнать что же умеет компьютер в области звука нужно только поинтересоваться и перед вами откроются возможности, о которых вы, может быть, даже не догадывались. И все это не так сложно, как может показаться на первый взгляд.

Некоторые факты и понятия, без которых тяжело обойтись.

В соответствии с теорией математика Фурье, звуковую волну можно представить в виде спектра входящих в нее частот (рис. 1).

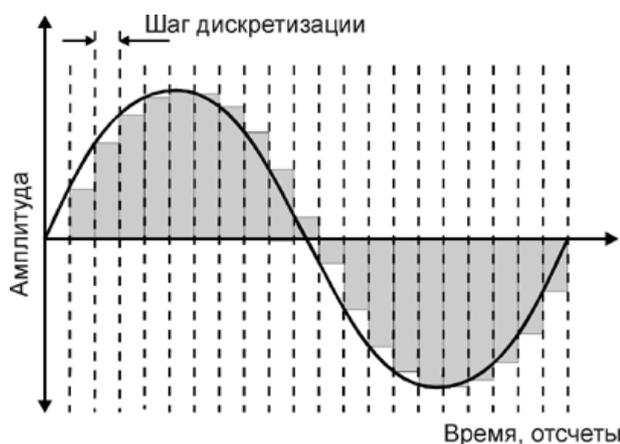


Рисунок 1.

Частотные составляющие спектра - это синусоидальные колебания (так называемые чистые тона), каждое из которых имеет свою собственную амплитуду и частоту. Таким образом, любое, даже самое сложное по форме колебание (например, человеческий голос), можно представить суммой простейших синусоидальных колебаний определенных частот и амплитуд. И наоборот, сгенерировав различные колебания и наложив их друг на друга (смикшировав, смешав), можно получить различные звуки.

Справка: человеческий слуховой аппарат/мозг способен различать частотные составляющие звука в пределах от 20 Гц до ~20 КГц (верхняя граница может колебаться в зависимости от возраста и других факторов). Кроме того, нижняя граница сильно колеблется в зависимости от интенсивности звучания.

1. Оцифровка звука и его хранение на цифровом носителе

«Обычный» аналоговый звук представляется в аналоговой аппаратуре непрерывным электрическим сигналом. Компьютер оперирует с данными в цифровом виде. Это означает, что и звук в компьютере представляется в цифровом виде. Как же происходит преобразование аналогового сигнала в цифровой?

Цифровой звук – это способ представления электрического сигнала посредством дискретных численных значений его амплитуды. Допустим, мы имеем аналоговую звуковую дорожку хорошего качества (говоря «хорошее качество» будем предполагать нешумную запись, содержащую спектральные составляющие из всего слышимого диапазона частот – приблизительно от 20 Гц до 20 КГц) и хотим «ввести» ее в компьютер (то есть оцифровать) без потери качества. Как этого добиться и как происходит оцифровка? Звуковая волна – это некая сложная функция, зависимость амплитуды звуковой волны от времени. Казалось бы, что раз это функция, то можно записать ее в компьютер «как есть», то есть описать математический вид функции и сохранить в памяти компьютера. Однако практически это невозможно, поскольку звуковые колебания нельзя представить аналитической формулой (как $y=x^2$, например). Остается один путь – описать функцию путем хранения ее дискретных значений в определенных точках. Иными словами, в каждой точке времени можно измерить значение амплитуды сигнала и записать в виде чисел. Однако и в этом методе есть свои недостатки, так как значения амплитуды сигнала мы не можем записывать с бесконечной точностью, и вынуждены их округлять. Говоря иначе, мы будем приближать эту функцию по двум координатным осям – амплитудной и временной (приближать в точках – значит, говоря простым языком, брать значения функции в точках и записывать их с конечной точностью). Таким образом, оцифровка сигнала включает в себя два процесса - процесс дискретизации (осуществление выборки) и процесс квантования. Процесс дискретизации - это процесс получения значений величин преобразуемого сигнала в определенные промежутки времени (рис. 2).

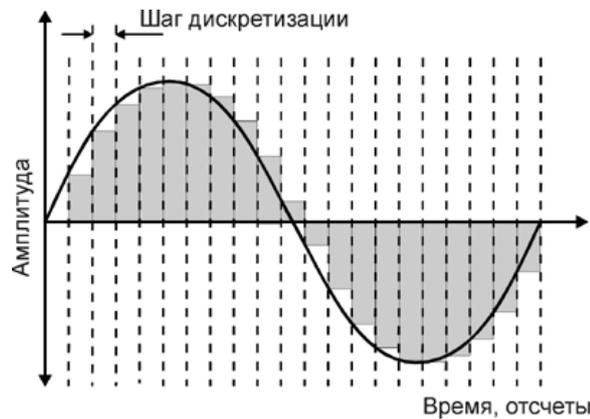


Рисунок 2

Квантование - процесс замены реальных значений сигнала приближенными с определенной точностью (рис. 3). Таким образом, оцифровка – это фиксация амплитуды сигнала через определенные промежутки времени и регистрация полученных значений амплитуды в виде округленных цифровых значений (так как значения амплитуды являются величиной непрерывной, нет возможности конечным числом записать точное значение амплитуды сигнала, именно поэтому прибегают к округлению). Записанные значения амплитуды сигнала называются отсчетами. Очевидно, что чем чаще мы будем делать замеры амплитуды (чем выше частота дискретизации) и чем меньше мы будем округлять полученные значения (чем больше уровней квантования), тем более точное представление сигнала в цифровой форме мы получим. Оцифрованный сигнал в виде набора последовательных значений амплитуды можно сохранить.

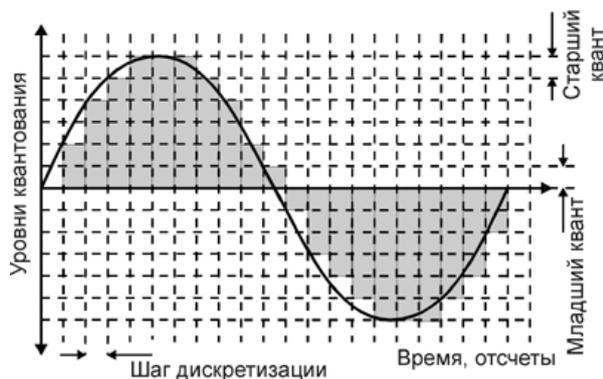


Рисунок 3

Теперь о практических проблемах. Во-первых, надо иметь в виду, что память компьютера не бесконечна, так что каждый раз при оцифровке необходимо находить какой-то компромисс между качеством (напрямую зависящим от использованных при оцифровке параметров) и занимаемым оцифрованным сигналом объемом.

Во-вторых, согласно теореме Котельникова частота дискретизации устанавливает верхнюю границу частот оцифрованного сигнала, а именно, максимальная частота спектральных составляющих равна половине частоты дискретизации сигнала. Попросту говоря, чтобы получить полную информацию о звуке в частотной полосе до 22050 Гц, необходима дискретизация с частотой не менее 44.1 КГц.

Существуют и другие проблемы и нюансы, связанные с оцифровкой звука. Не сильно углубляясь в подробности отметим, что в «цифровом звуке» из-за дискретности информации об амплитуде оригинального сигнала появляются различные шумы и искажения (под фразой «в цифровом звуке есть такие-то частоты и шумы» подразумевается, что когда этот звук будет преобразован обратно из цифрового вида в аналоговый, то в его звучании будут присутствовать упомянутые частоты и шумы). Так, например, джиттер (jitter) – шум, появляющийся в результате того, что осуществление выборки сигнала при дискретизации происходит не через абсолютно равные промежутки времени, а с какими-то отклонениями. То есть, если, ска-

жем, дискретизация проводится с частотой 44.1 КГц, то отсчеты берутся не точно каждые 1/44100 секунды, а то немного раньше, то немного позднее. А так как входной сигнал постоянно меняется, то такая ошибка приводит к «захвату» не совсем верного уровня сигнала. В результате во время проигрывания оцифрованного сигнала может ощущаться некоторое дрожание и искажения. Появление джиттера является результатом не абсолютной стабильности аналогово-цифровых преобразователей. Для борьбы с этим явлением применяют высокостабильные тактовые генераторы. Еще одной неприятностью является шум дробления. Как мы говорили, при квантовании амплитуды сигнала происходит ее округление до ближайшего уровня. Такая погрешность вызывает ощущение «грязного» звучания.

Небольшая справка: стандартные параметры записи аудио компакт-дисков следующие: частота дискретизации - 44.1 КГц, уровень квантования – 16 бит. Такие параметры соответствуют 65536 (2¹⁶) уровням квантования амплитуды при взятии ее значений 44100 раз в секунду.

На практике, процесс оцифровки (дискретизация и квантование сигнала) остается невидимым для пользователя - всю черновую работу делают разнообразные программы, которые дают соответствующие команды драйверу (управляющая подпрограмма операционной системы) звуковой карты. Любая программа (будь то встроенный в Windows Recorder или мощный звуковой редактор), способная осуществлять запись аналогового сигнала в компьютер, так или иначе оцифровывает сигнал с определенными параметрами, которые могут оказаться важными в последующей работе с записанным звуком, и именно по этой причине важно понять как происходит процесс оцифровки и какие факторы влияют на ее результаты.

2. Преобразование звука из цифрового вида в аналоговый

Как после оцифровки прослушивать звук? То есть, как преобразовывать его обратно из цифрового вида в аналоговый?

Для преобразования дискретизованного сигнала в аналоговый вид, пригодный для обработки аналоговыми устройствами (усилителями и фильтрами) и последующего воспроизведения через акустические системы, служит цифроаналоговый преобразователь (ЦАП). Процесс преобразования представляет собой обратный процесс дискретизации: имея информацию о величине отсчетов (амплитуды сигнала) и беря определенное количество отсчетов в единицу времени, путем интерполирования происходит восстановление исходного сигнала (рис. 4).

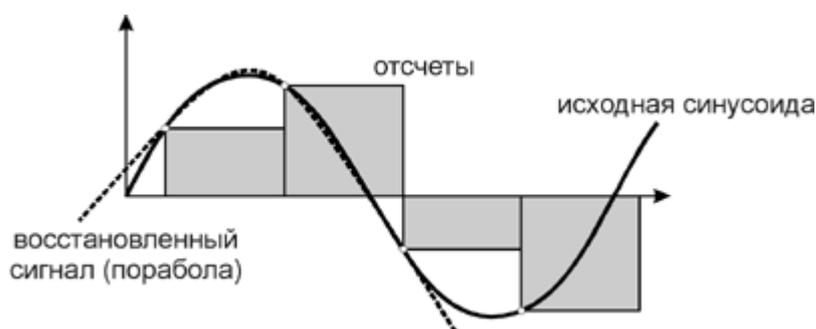


Рисунок 4

Еще совсем недавно воспроизведение звука в домашних компьютерах было проблемой, так как компьютеры не оснащались специальными ЦАП. Сначала в качестве простейшего звукового устройства в компьютере использовался встроенный динамик (PC speaker). Вообще говоря, этот динамик до сих пор имеется почти во всех PC, но никто уже не помнит как его «раскачать», чтобы он заиграл. Если вкратце, то этот динамик присоединен к порту на материнской плате, у которого есть два положения – 1 и 0. Так вот, если этот порт быстро включать и выключать, то из динамика можно извлечь более-менее правдоподобные звуки. Воспроизведение различных частот достигается за счет того, что диффузор динамика обладает конечной реакцией и не способен мгновенно перескакивать с места на место, таким образом он «плавно раскачивается» вследствие скачкообразного изменения напряжения на нем. И если колебать его с разной скоростью, то можно получить колебания воздуха на разных частотах. Естественной альтернативой динамику стал так называемый Covox – это про-

стейший ЦАП, выполненный на нескольких подобранных сопротивлениях (или готовой микросхеме), которые обеспечивают перевод цифрового представления сигнала в аналоговый – то есть в реальные значения амплитуды. Sovox прост в изготовлении и поэтому он пользовался успехом у любителей вплоть до того времени, когда звуковая карта стала доступной всем.

В современном компьютере звук воспроизводится и записывается с помощью звуковой карты, подключаемой либо встроенной в материнскую плату компьютера. Задача звуковой карты в компьютере – ввод и вывод аудио. Практически это означает, что звуковая карта является тем преобразователем, который переводит аналоговый звук в цифровой и обратно. Если описывать упрощенно, то работа звуковой карты может быть пояснена следующим образом. Предположим, что на вход звуковой карты подан аналоговый сигнал и карта включена (программно) в режим. Сначала входной аналоговый сигнал попадает в аналоговый микшер, который занимается смешением сигналов и регулировкой громкости и баланса. Микшер необходим, в частности, для предоставления возможности пользователю управлять уровнями. Затем отрегулированный и сбалансированный сигнал попадает в аналогово-цифровой преобразователь, где сигнал дискретизируется и квантуется, в результате чего в компьютер по шине данных направляется бит-поток, который и представляет собой оцифрованный аудио сигнал. Вывод аудио информации почти аналогичен вводу, только происходит в обратную сторону. Поток данных, направленный в звуковую карту, преодолевает цифро-аналоговый преобразователь, который образует из чисел, описывающих амплитуду сигнала, электрический сигнал; полученный аналоговый сигнал может быть пропущен через любые аналоговые тракты для дальнейших преобразований, в том числе и для воспроизведения. Надо отметить, что если звуковая карта оборудована интерфейсом для обмена цифровыми данными, то при работе с цифровым аудио никакие аналоговые блоки карты не задействуются.

3. Способы хранения цифрового звука

Для хранения цифрового звука существует много различных способов. Как мы говорили, оцифрованный звук является набором значений амплитуды сигнала, взятых через определенные промежутки времени. Таким образом, во-первых, блок оцифрованной аудио информации можно записать в файл «как есть», то есть последовательностью чисел (значений амплитуды). В этом случае существуют два способа хранения информации.

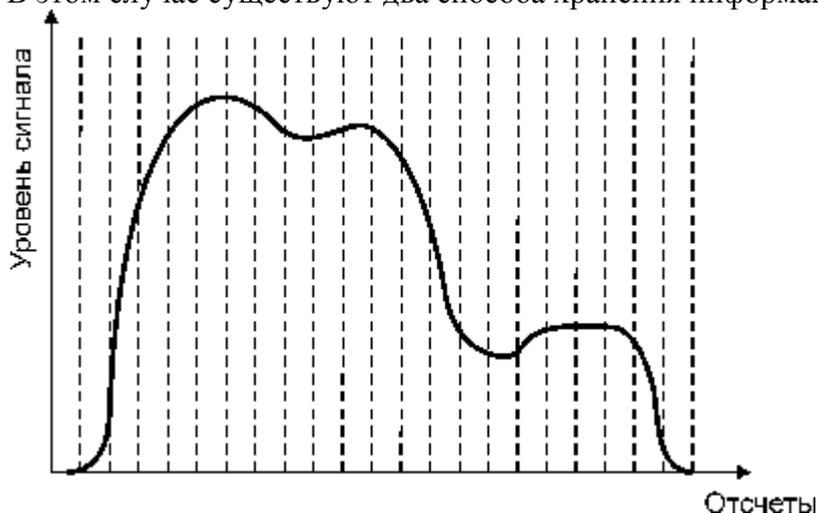


Рисунок 5

Первый (рис. 5) - PCM (Pulse Code Modulation - импульсно-кодовая модуляция) - способ цифрового кодирования сигнала при помощи записи абсолютных значений амплитуд (бывают знаковое или беззнаковое представления). Именно в таком виде записаны данные на всех аудио CD. Второй способ (рис. 6) - ADPCM (Adaptive Delta PCM - адаптивная относительная импульсно-кодовая модуляция) – запись значений сигнала не в абсолютных, а в относительных изменениях амплитуд (приращениях).

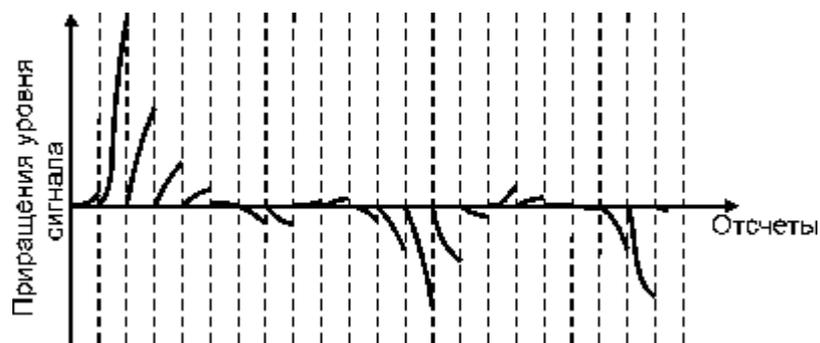


Рисунок 6

Во-вторых, можно сжать или упростить данные так, чтобы они занимали меньший объем памяти, нежели будучи записанными «как есть». Тут тоже имеются два пути.

Кодирование данных без потерь (lossless coding) - это способ кодирования аудио, который позволяет осуществлять стопроцентное восстановление данных из сжатого потока. К такому способу уплотнения данных прибегают в тех случаях, когда сохранение оригинального качества данных критично. Например, после сведения звука в студии звукозаписи, данные необходимо сохранить в архиве в оригинальном качестве для возможного последующего использования. Существующие сегодня алгоритмы кодирования без потерь (например, Monkeys Audio) позволяют сократить занимаемый данными объем на 20-50%, но при этом обеспечить стопроцентное восстановление оригинальных данных из полученных после сжатия. Подобные кодеры – это своего рода архиваторы данных (как ZIP, RAR и другие), только предназначенные для сжатия именно аудио.

Имеется и второй путь кодирования, на котором мы остановимся чуть подробнее, – кодирование данных с потерями (lossy coding). Цель такого кодирования - любыми способами добиться схожести звучания восстановленного сигнала с оригиналом при как можно меньшем объеме упакованных данных. Это достигается путем использования различных алгоритмов «упрощающих» оригинальный сигнал (выкидывая из него «ненужные» слабослышимые детали), что приводит к тому, что декодированный сигнал фактически перестает быть идентичным оригиналу, а лишь похоже звучит. Методов сжатия, а также программ, реализующих эти методы, существует много. Наиболее известными являются MPEG-1 Layer I, II, III (последним является всем известный MP3), MPEG-2 AAC (advanced audio coding), Ogg Vorbis, Windows Media Audio (WMA), TwinVQ (VQF), MPEGPlus, TAC, и прочие. В среднем, коэффициент сжатия, обеспечиваемый такими кодерами, находится в пределах 10-14 (раз). Надо особо подчеркнуть, что в основе всех lossy-кодеров лежит использование так называемой психоакустической модели, которая как раз и занимается «упрощением» оригинального сигнала. Говоря точнее, механизм подобных кодеров выполняет анализ кодируемого сигнала, в процессе которого определяются участки сигнала, в определенных частотных областях которых имеются неслышимые человеческому уху нюансы (замаскированные или неслышимые частоты), после чего происходит их удаление из оригинального сигнала. Таким образом, степень сжатия оригинального сигнала зависит от степени его «упрощения»; сильное сжатие достигается путем «агрессивного упрощения» (когда кодер «считает» ненужными множественные нюансы), такое сжатие, естественно, приводит к сильной деградации качества, поскольку удалению могут подлежать не только незаметные, но и значимые детали звучания.

Как мы сказали, современных lossy-кодеров существует достаточно много. Наиболее распространенный формат – MPEG-1 Layer III (всем известный MP3). Формат завоевал свою популярность совершенно заслуженно – это был первый распространенный кодек подобного рода, который достиг столь высокого уровня компрессии при отличном качестве звучания. Сегодня этому кодеку имеется множество альтернатив, выбор остается за пользователем. К сожалению, рамки статьи не позволяют привести здесь тестирования и сравнения существующих кодеков, однако авторы статьи позволяют себе привести некоторую информацию, полезную при выборе кодека. Итак, преимущества MP3 – широкая распространенность и достаточно высокое качество кодирования, которое объективно улучшается благодаря разра-

боткам различных кодеров MP3 энтузиастами (например, кодек Lame). Мощная альтернатива MP3 – кодек Microsoft Windows Media Audio (Файлы .WMA и .ASF). По различным тестам этот кодек показывает себя от «как MP3» до «заметно хуже MP3» на средних битрейтах, и, чаще, «лучше MP3» на низких битрейтах. Ogg Vorbis (файлы .OGG) – совершенно свободный от лицензирования кодек, создаваемый независимыми разработчиками. Чаще всего ведет себя лучше MP3, недостатком является лишь малая распространенность, что может стать критическим аргументом при выборе кодека для длительного хранения аудио. Вспомним и еще молодой кодек MP3 Pro, анонсированный в июле 2001 года компанией Coding Technologies совместно с Thomson Multimedia. Кодек является продолжением, или, точнее, развитием старого MP3 – он совместим с MP3 назад (полностью) и вперед (частично). За счет использования новой технологии SBR (Spectral Band Replication), кодек ведет себя заметно лучше других форматов на низких битрейтах, однако качество кодирования на средних и высоких битрейтах чаще уступает качеству почти всех описанных кодеков. Таким образом, MP3 Pro пригоден больше для ведения аудио трансляций в Internet, а также для создания превью песен и музыки.

Говоря о способах хранения звука в цифровом виде нельзя не вспомнить и о носителях данных. Всем привычный аудио компакт-диск, появившийся в начале 80-х годов, широкое распространение получил именно в последние годы (что связано с сильным удешевлением носителя и приводов). А до этого носителями цифровых данных являлись кассеты с магнитной лентой, но не обычные, а специально предназначенные для так называемых DAT-магнитофонов. Ничего примечательного – магнитофоны как магнитофоны, однако цена на них всегда была высокой, и такое удовольствие было не всем «по зубам». Эти магнитофоны использовались, в основном, в студиях звукозаписи. Преимущество таких магнитофонов было в том, что, не смотря на использование привычных носителей, данные на них хранились в цифровом виде и практически никаких потерь при чтении/записи на них не было (что очень важно при студийной обработке и хранении звука). Сегодня появилось большое количество различных носителей данных, кроме привычных всем компакт дисков. Носители совершенствуются и с каждым годом становятся более доступными и компактными. Это открывает большие возможности в области создания мобильных аудио проигрывателей. Уже сегодня продается огромное количество различных моделей переносных цифровых плееров. И, можно предположить, что это еще далеко не пик развития такого рода техники.

4. Преимущества и недостатки цифрового звука

С точки зрения обычного пользователя выгоды много - компактность современных носителей информации позволяет ему, например, перевести все диски и пластинки из своей коллекции в цифровое представление и сохранить на долгие годы на небольшом трехдюймовом винчестере или на десятке-другом компакт дисков; можно воспользоваться специальным программным обеспечением и хорошенько «почистить» старые записи с бобин и пластинок, удалив из их звучания шумы и треск; можно также не просто скорректировать звучание, но и приукрасить его, добавить сочности, объемности, восстановить частоты. Помимо перечисленных манипуляций со звуком в домашних условиях, Интернет тоже приходит на помощь аудио-любителю. Например, сеть позволяет людям обмениваться музыкой, прослушивать сотни тысяч различных Интернет-радио станций, а также демонстрировать свое звуковое творчество публике, и для этого нужен всего лишь компьютер и Интернет. И, наконец, в последнее время появилась огромная масса различной портативной цифровой аудио аппаратуры, возможности даже самого среднего представителя которой зачастую позволяют с легкостью взять с собой в дорогу коллекцию музыки, равную по длительности звучания десяткам часов.

С точки зрения профессионала цифровой звук открывает поистине необъятные возможности. Если раньше звуковые и радио студии размещались на нескольких десятках квадратных метров, то теперь их может заменить хороший компьютер, который по возможностям превосходит десять таких студий вместе взятых, а по стоимости оказывается многократно дешевле одной. Это снимает многие финансовые барьеры и делает звукозапись более до-

ступной и профессионалу и простому любителю. Современное программное обеспечение позволяет делать со звуком все что угодно. Раньше различные эффекты звучания достигались с помощью хитроумных приспособлений, которые не всегда являли собой верх технической мысли или же были просто устройствами кустарного изготовления. Сегодня, самые сложные и просто невообразимые раньше эффекты достигаются путем нажатия пары кнопок. Конечно, вышесказанное несколько утрировано и компьютер не заменяет человека – звукооператора, режиссера или монтажера, однако с уверенностью можно сказать, что компактность, мобильность, колоссальная мощность и обеспечиваемое качество современной цифровой техники, предназначенной для обработки звука, уже сегодня почти полностью вытеснило из студий старую аналоговую аппаратуру.

Конечно, цифровая техника тоже имеет свои недостатки. Многие (профессионалы и любители) отмечают, что аналоговый звук слушался живее. И это не просто дань прошлому. Как мы сказали выше, процесс оцифровки вносит определенную погрешность в звучание, кроме того, различная усиливающая цифровая аппаратура привносит так называемые «транзисторные шумы» и другие специфические искажения. Термину «транзисторный шум», пожалуй, нет точного определения, но можно сказать, что это хаотичные колебания в области высоких частот. Не смотря на то, что слуховой аппарат человека способен воспринимать частоты до 20 кГц, похоже, все-таки, человеческий мозг улавливает и более высокие частоты. И именно на подсознательном уровне человек все же ощущает аналоговое звучание чище, чем цифровое.

Впрочем, у цифрового представления данных есть одно неоспоримое и очень важное преимущество – при сохранном носителе данные на нем не искажаются с течением времени. Если магнитная лента со временем размагничивается и качество записи теряется, если пластинка царапается и к звучанию прибавляются щелчки и треск, то компакт-диск / винчестер / электронная память либо читается (в случае сохранности), либо нет, а эффект старения отсутствует. Важно отметить, мы не говорим здесь об Audio CD (CD-DA – стандарт, устанавливающий параметры и формат записи на аудио компакт диски) так как не смотря на то, что это носитель цифровой информации, эффект старения его, все же, не минует. Это связано с особенностями хранения и считывания аудио данных с Audio CD. Информация на всех типах компакт-дисков хранится покадрово и каждый кадр имеет заголовок, по которому его возможно идентифицировать. Однако различные типы CD имеют различную структуру и используют различные методы маркировки кадров. Поскольку компьютерные приводы CD-ROM рассчитаны на чтение в основном Data-CD (надо сказать, что существуют различные разновидности стандарта Data-CD, каждый из которых дополняет основной стандарт CD-DA), они часто не способны правильно «ориентироваться» на Audio CD, где способ маркировки кадров отличен от Data-CD (на аудио CD кадры не имеют специального заголовка и для определения смещения каждого кадра необходимо следить за информацией в кадре). Это означает, что если при чтении Data-CD привод легко «ориентируется» на диске и никогда не перепутает кадры, то при чтении с аудио компакт диска привод не может ориентироваться четко, что при появлении, скажем, царапины или пыли может привести к чтению неправильного кадра и, как следствие, скачку или треску звучания. Эта же проблема (неспособность большинства приводов правильно позиционироваться на CD-DA) является причиной еще одного неприятного эффекта: копирование информации с Audio CD вызывает проблемы даже при работе с полностью сохраненными дисками вследствие того, что правильное «ориентирование на диске» полностью зависит от считывающего привода и не может быть четко проконтролировано программным путем.

Повсеместное распространение и дальнейшее развитие уже упомянутых lossy-кодеров аудио (MP3, AAC и других) открыло широчайшие возможности распространения и хранения аудио. Современные каналы связи уже давно позволяют пересылать большие массивы данных за сравнительно небольшое время, однако самой медленной остается передача данных между конечным пользователем и поставщиком услуг связи. Телефонные линии, по которым пользователи в большинстве своем связываются с Интернетом, не позволяют осуществлять

быструю передачу данных. Нечего и говорить, что такие объемы данных, какие занимает несжатая аудио и видео информация, передавать по привычным каналам связи придется очень долго. Однако появление lossy-кодеров, обеспечивающих десяти-пятнадцатикратное сжатие, превратило передачу и обмен аудио данными в повседневное занятие каждого пользователя Интернета и сняло все преграды, образованные слабыми каналами связи. Касательно этого нужно сказать, что развивающаяся сегодня семимильными шагами цифровая мобильная связь во многом обязана именно lossy-кодированию. Дело в том, что протоколы передачи аудио по каналам мобильной связи работают на приблизительно тех же принципах, что и известные всем музыкальные кодеры. Поэтому дальнейшее развитие в области кодирования аудио неизменно ведет к уменьшению стоимости передачи данных в мобильных системах, от чего конечный пользователь только выигрывает: дешевеет связь, появляются новые возможности, продлевается время работы батарей мобильных устройств и т.д. Не в меньшей степени lossy-кодирование помогает экономить деньги на покупке дисков с любимыми песнями – сегодня стоит только зайти в Интернет и там можно найти почти любую интересующую песню. Безусловно, такое положение вещей давно «мозолит глаза» звукозаписывающим компаниям – у них под носом люди вместо покупки дисков обмениваются песнями прямо через Интернет, что превращает некогда золотое дно в малоприбыльный бизнес, но это уже вопрос этики и финансов. Одно можно сказать с уверенностью: с таким положением вещей уже ничего нельзя поделать и бум обмена музыкой через Интернет, порожденный именно появлением lossy-кодеров, уже ничем не остановить. А это только на руку рядовому пользователю.

5. К вопросу об обработке звука

Под обработкой звука следует понимать различные преобразования звуковой информации с целью изменения каких-то характеристик звучания. К обработке звука относятся способы создания различных звуковых эффектов, фильтрация, а также методы очистки звука от нежелательных шумов, изменения тембра и т.д. Все это огромное множество преобразований сводится, в конечном счете, к следующим основным типам:

1. Амплитудные преобразования. Выполняются над амплитудой сигнала и приводят к ее усилению/ослаблению или изменению по какому-либо закону на определенных участках сигнала.

2. Частотные преобразования. Выполняются над частотными составляющими звука: сигнал представляется в виде спектра частот через определенные промежутки времени, производится обработка необходимых частотных составляющих, например, фильтрация, и обратное «сворачивание» сигнала из спектра в волну.

3. Фазовые преобразования. Сдвиг фазы сигнала тем или иным способом; например, такие преобразования стерео сигнала, позволяют реализовать эффект вращения или «объемности» звука.

4. Временные преобразования. Реализуются путем наложения, растягивания/сжатия сигналов; позволяют создать, например, эффекты эха или хора, а также повлиять на пространственные характеристики звука.

Обсуждение каждого из названных типов преобразований может стать целым научным трудом. Стоит привести несколько практических примеров использования указанных видов преобразований при создании реальных звуковых эффектов:

Echo (эхо) Реализуется с помощью временных преобразований. Фактически для получения эха необходимо на оригинальный входной сигнал наложить его задержанную во времени копию. Для того, чтобы человеческое ухо воспринимало вторую копию сигнала как повторение, а не как отзвук основного сигнала, необходимо время задержки установить равным примерно 50 мс. На основной сигнал можно наложить не одну его копию, а несколько, что позволит на выходе получить эффект многократного повторения звука (многоголосного эха). Чтобы эхо казалось затухающим, необходимо на исходный сигнал накладывать не просто задержанные копии сигнала, а приглушенные по амплитуде.

Reverberation (повторение, отражение). Эффект заключается в придании звучанию объемности, характерной для большого зала, где каждый звук порождает соответствующий,

медленно угасающий отзвук. Практически, с помощью реверберации можно «оживить», например, фонограмму, сделанную в заглушенном помещении. От эффекта «эхо» реверберация отличается тем, что на входной сигнал накладывается задержанный во времени выходной сигнал, а не задержанная копия входного. Иными словами, блок реверберации упрощенно представляет собой петлю, где выход блока подключен к его входу, таким образом уже обработанный сигнал каждый цикл снова подается на вход смешиваясь с оригинальным сигналом.

Chorus (хор). В результате его применения звучание сигнала превращается как бы в звучание хора или в одновременное звучание нескольких инструментов. Схема получения такого эффекта аналогична схеме создания эффекта эха с той лишь разницей, что задержанные копии входного сигнала подвергаются слабой частотной модуляции (в среднем от 0.1 до 5 Гц) перед смешиванием со входным сигналом. Увеличение количества голосов в хоре достигается путем добавления копий сигнала с различными временами задержки.

Безусловно, как и во всех других областях, в обработке сигналов также имеются проблемы, которые являются своего рода камнем преткновения. Так, например, при разложении сигналов в спектр частот существует принцип неопределенности, который невозможно преодолеть. Принцип гласит, что нельзя получить точную спектральную картину сигнала в конкретный момент времени: либо для получения более точной спектральной картины нужно проанализировать большой временной участок сигнала, либо, если нас интересует больше время, когда происходило то или иное изменение спектра, нужно пожертвовать точностью самого спектра. Иными словами нельзя получить точный спектр сигнала в точке - точный спектр для большого участка сигнала, либо очень приблизительный спектр, но для короткого участка.

Механизмы для обработки сигналов существуют как в программном, так и в аппаратном исполнении (так называемые эффект-процессоры). Например, вокодеры и гитарные процессоры, хорусы и ревербераторы существуют в виде аппаратуры, а также в виде программ.

Практическую обработку сигналов можно разделить на два типа: обработка «на лету» и пост-обработка. Обработка «на лету» подразумевает мгновенное преобразование сигнала (то есть с возможностью осуществлять вывод обработанного сигнала почти одновременно с его вводом). Простой пример – гитарные «примочки» или реверберация во время живого исполнения на сцене. Такая обработка происходит мгновенно, то есть, скажем, исполнитель поет в микрофон, а эффект-процессор преобразует его голос и слушатель слышит уже обработанный вариант голоса. Пост-обработка – это обработка уже записанного сигнала. Скорость такой обработки может быть сильно ниже скорости воспроизведения. Такая обработка преследует те же цели, то есть придание звуку определенного характера, либо изменение характеристик, однако применяется на стадии мастеринга или подготовки звука к тиражированию, когда не требуется спешка, а важнее качество и скрупулезная проработка всех нюансов звучания. Существует множество различных операций над звуком, которые вследствие недостаточной производительности сегодняшних процессоров нельзя реализовать «на лету», поэтому такие преобразования проводят лишь в пост-режиме.

Обработка сигнала – это сложная и, главное, ресурсоемкая процедура. Она сравнительно недавно стала проводиться в цифровых устройствах – раньше различные эффекты звучания и другие достигались путем обработки звука в аналоговых приборах. В аналоговой аппаратуре звук в виде электрических колебаний проходит через различные тракты (блоки электрических элементов), чем достигается изменение фазы, спектра и амплитуды сигнала. Однако такой способ обработки имеет массу недостатков. Во-первых, страдает качество обработки, ведь каждый аналоговый элемент имеет свою погрешность, а несколько десятков элементов могут критически повлиять на точность и качество желаемого результата. А во-вторых, и это, пожалуй, самое главное, почти каждый отдельный эффект достигается путем использования отдельного устройства, когда каждое такое устройство может стоить очень дорого. Возможность же использования цифровых устройств имеет неоспоримые преиму-

щества. Качество обработки сигналов в них намного меньше зависит от качества аппаратуры, главное – это качественно оцифровать звук и иметь возможность качественно его воспроизводить, и тогда качество обработки ложится уже только на программный механизм. Кроме того, для различных манипуляций со звуком не требуется постоянная смена оборудования. И, самое главное, поскольку обработка ведется программным путем, для нее открываются просто невероятные возможности, которые ограничены лишь мощностью компьютеров (а она увеличивается с каждым днем) и фантазией человека. Однако, (по крайней мере сегодня) здесь имеются и свои неприятности. Так, например, часто, даже для осуществления несложной обработки сигнала необходимо осуществить его разложение в спектр частот. В этом случае обработка сигнала на лету может быть затруднена именно из-за ресурсоемкости этапа разложения. Поэтому преобразования, требующие спектрального разложения, выполняются чаще в пост-режиме.

6. Аппаратура

Немаловажная часть разговора о звуке связана с аппаратурой. Существует много различных устройств для обработки и ввода/вывода звука. Касательно обычного персонального компьютера следует подробнее остановиться на звуковых картах. Звуковые карты принято делить на звуковые, музыкальные и звукомузыкальные. По конструкции же все звуковые платы можно разделить на две группы: основные (устанавливаемые на материнской плате компьютера и обеспечивающие ввод и вывод аудио данных) и дочерние (имеют принципиальное конструктивное отличие от основных плат - они чаще всего подключаются к специальному разъему, расположенному на основной плате). Дочерние платы служат чаще всего для обеспечения или расширения возможностей MIDI-синтезатора.

Звукомузыкальные и звуковые платы выполняются в виде устройств, вставляемых в слот материнской платы (либо уже встроены в нее изначально). Визуально они имеют обычно два аналоговых входа - линейный и микрофонный, и несколько аналоговых выходов: линейные выходы и выход для наушников. В последнее время карты стали оснащаться также и цифровым входом и выходом, обеспечивающим передачу аудио между цифровыми устройствами. Аналоговые входы и выходы обычно имеют разъемы, аналогичные разъемам головных наушников (1/8"). Вообще, входов у звуковой платы немного больше, чем два: аналоговые CD, MIDI и другие входы. Они, в отличие от микрофонного и линейного входов, расположены не на задней панели звуковой платы, а на самой плате; могут иметься и другие входы, например, для подключения голосового модема. Цифровые входы и выходы обычно выполнены в виде интерфейса S/PDIF (интерфейс цифровой передачи сигналов) с соответствующим разъемом (S/PDIF – сокращение от Sony/Panasonic Digital Interface - цифровой интерфейс Sony/Panasonic). S/PDIF - это «бытовой» вариант более сложного профессионального стандарта AES/EBU (Audio Engineering Society / European Broadcast Union). Сигнал S/PDIF используется для цифровой передачи (кодирования) 16-разрядных стерео данных с любой частотой дискретизации. Помимо перечисленного, на звукомузыкальных платах имеется MIDI-интерфейс с разъемами для подключения MIDI-устройств и джойстиков, а также для подсоединения дочерней музыкальной карты (хотя в последнее время возможность подключения последней становится редкостью). Некоторые модели звуковых карт для удобства пользователя оснащаются фронтальной панелью, устанавливаемой на лицевой стороне системного блока компьютера, на которой размещаются разъемы, соединенные с различными входами и выходами звуковой карты.

Определим несколько основных блоков, из которых состоят звуковые и звукомузыкальные платы.

1. Блок цифровой обработки сигналов (кодек). В этом блоке осуществляются аналого-цифровые и цифро-аналоговые преобразования (АЦП и ЦАП). От этого блока зависят такие характеристики карты, как максимальная частота дискретизации при записи и воспроизведении сигнала, максимальный уровень квантования и максимальное количество обрабатываемых каналов (моно или стерео). В немалой степени от качества и сложности составляющих этого блока зависят и шумовые характеристики.

2. Блок синтезатора. Присутствует в музыкальных картах. Выполняется на основе либо FM-, либо WT-синтеза, либо на обоих сразу. Может работать как под управлением собственного процессора, так и под управлением специального драйвера.

3. Интерфейсный блок. Обеспечивает передачу данных по различным интерфейсам (например, S/PDIF). У чисто звуковой карты этот блок чаще отсутствует.

4. Микшерный блок. В звуковых платах микшерный блок обеспечивает регулировку: уровней сигналов с линейных входов; уровней с MIDI входа и входа цифрового звука; уровня общего сигнала; панорамирования; тембра.

Рассмотрим важнейшие параметры, характеризующие звуковые и звукомзыкальные платы. Наиболее важными характеристиками являются: максимальная частота дискретизации (sampling rate) в режиме записи и в режиме воспроизведения, максимальный уровень квантования или разрядность (max. quantization level) в режиме записи и воспроизведения. Кроме того, так как звукомзыкальные платы имеют еще и синтезатор, то к их характеристикам относят и параметры установленного синтезатора. Естественно, чем с большим уровнем квантования карта способна кодировать сигналы, тем большее качество сигнала при этом достигается. Все современные модели звуковых карт способны кодировать сигнал с уровнем 16 бит. Одной из важных характеристик является возможность одновременного воспроизведения и записи звуковых потоков. Особенность карты одновременно воспроизводить и записывать называют полнодуплексной (full duplex). Есть еще одна характеристика, которая зачастую играет решающую роль при покупке звуковой карты - отношение сигнал/шум (Signal/Noise Ratio, S/N). Этот показатель влияет на чистоту записи и воспроизведения сигнала. Отношение сигнал/шум – это отношение мощности сигнала к мощности шума на выходе устройства, этот показатель принято измерять в дБ. Хорошим можно считать отношение 80-85 дБ; идеальным – 95-100 дБ. Однако нужно учитывать, что на качество воспроизведения и записи сильно влияют наводки (помехи) со стороны других компонент компьютера (блока питания и проч.). В результате этого отношение сигнал/шум может изменяться в худшую сторону. На практике методов борьбы с этим существует достаточно много. Некоторые предлагают заземлить компьютер. Другие, дабы как можно более тщательно уберечь звуковую карту от наводок, «выносят» ее за пределы корпуса компьютера. Однако полностью уберечься от наводок очень тяжело, так как даже элементы самой карты создают наводки друг на друга. С этим тоже пытаются бороться и для этого экранируют каждый элемент на плате. Но сколько бы усилий не прилагалось к решению этой проблемы, полностью исключить влияние внешних помех невозможно.

Еще одна не менее важная характеристика – коэффициент нелинейных искажений или Total Harmonic Distortion, THD. Этот показатель также критическим образом влияет на чистоту звучания. Коэффициент нелинейных искажений измеряется в процентах: 1% - «грязное» звучание; 0.1% - нормальное звучание; 0.01% - чистое звучание класса Hi-Fi; 0.002% - звучание класса Hi-Fi – Hi End. Нелинейные искажения – результат неточности в восстановлении сигнала из цифрового вида в аналоговый. Упрощенно, процесс измерения этого коэффициента проводится следующим образом. На вход звуковой карты подается чистый синусоидальный сигнал. На выходе устройства снимается сигнал, спектр которого представляет собой сумму синусоидальных сигналов (сумма исходной синусоиды и ее гармоник). Затем по специальной формуле рассчитывается количественное соотношение исходного сигнала и его гармоник, полученных на выходе устройства. Это количественное соотношение и есть коэффициент нелинейных искажений (THD).

Что такое MIDI-синтезатор? Термин «синтезатор» обычно используется применительно к электронному музыкальному инструменту, в котором звук создается и обрабатывается, меняя свою окраску и характеристики. Естественно, название этого устройства пошло от его основного предназначения – синтеза звука. Основных методов синтеза звука существует всего два: FM (Frequency modulation – частотная модуляция) и WT (Wave Table – таблично-волновой). Поскольку мы не можем здесь подробно останавливаться на их рассмотрении, опишем лишь основную идею методов. В основе FM-синтеза лежит идея, что любое да-

же самое сложное колебание является по сути суммой простейших синусоидальных. Таким образом, можно наложить друг на друга сигналы от конечного числа генераторов синусоид и путем изменения частот синусоид получать звуки, похожие на настоящие. Таблично-волновой синтез основывается на другом принципе. Синтез звука при использовании такого метода достигается за счет манипуляций над заранее записанными (оцифрованными) звуками реальных музыкальных инструментов. Эти звуки (они называются сэмплами) хранятся в постоянной памяти синтезатора.

MIDI-синтезатор – это синтезатор, отвечающий требованиям стандарта, о котором мы сейчас поговорим. MIDI – это общепринятая спецификация, связанная с организацией цифрового интерфейса для музыкальных устройств, включающая в себя стандарт на аппаратную и программную части.

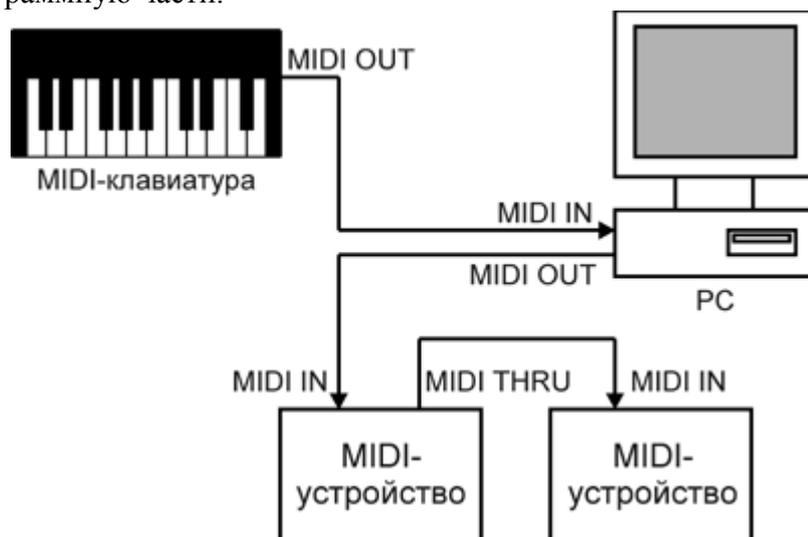


Рисунок. 7

Эта спецификация предназначена для организации локальной сети электронных инструментов (рис. 7). К MIDI-устройствам относятся различные аппаратные и музыкальные инструменты, отвечающие требованиям MIDI. Таким образом, MIDI-синтезатор – это музыкальный инструмент, предназначенный обычно для синтеза звука и музыки, а также удовлетворяющий спецификации MIDI. Давайте разберемся кратко, почему выделен отдельный класс устройств, названный MIDI.

Дело в том, что осуществление программной обработки звука часто сопряжено с неудобствами, обусловленными различными техническими особенностями этого процесса. Даже возложив операции по обработке звука на звуковую карту или любую другую аппаратуру, остается множество различных проблем. Во-первых, зачастую желательно пользоваться аппаратным синтезом звучания музыкальных инструментов (как минимум потому, что компьютер – это слишком общий инструмент, часто необходим просто аппаратный синтезатор звуков и музыки, не более). Во-вторых, программная обработка звука часто сопровождается временными задержками, в то время как при концертной работе необходимо мгновенное получение обработанного сигнала. По этим и другим причинам и прибегают к использованию специальной аппаратуры для обработки, а не компьютеров со специальными программами. Однако при использовании аппаратуры возникает необходимость в едином стандарте, который позволил бы соединять устройства друг с другом и комбинировать их. Эти предпосылки и заставили в 1982 году несколько ведущих в области музыкального оборудования компаний утвердить первый MIDI-стандарт, который впоследствии получил продолжение и развивается по сей день. Что же в конечном счете представляет собой MIDI-интерфейс и устройства в него входящие с точки зрения персонального компьютера?

Аппаратно - это установленные на звуковой карте: синтезатор различных звуков и музыкальных инструментов, микропроцессор, контролирующий и управляющий работу MIDI-устройств, а также различные стандартизованные разъемы и шнуры для подключения дополнительных устройств.

Программно - это протокол MIDI, представляющий собой набор сообщений (команд), которые описывают различные функции системы MIDI и с помощью которых осуществляется связь (обмен информацией) между устройствами MIDI. Сообщения можно рассматривать как средство удаленного управления.

Рамки данной статьи не позволяют нам углубляться в частности описания MIDI, следует отметить однако, что в отношении синтезаторов звука MIDI устанавливает строгие требования к их возможностям, примененным в них способам синтеза звука, а также к управляющим параметрам синтеза. Кроме того, для того, чтобы музыка созданная на одном синтезаторе могла бы быть легко перенесена и успешно воспроизведена на другом, были установлены несколько стандартов на соответствие инструментов (голосов) и их параметров в различных синтезаторах: стандарт General MIDI (GM), General Synth (GS) и eXtended General (XG). Базисным стандартом является GM, остальные два являются его логическими продолжениями и расширениями.

В качестве практического примера устройства MIDI, можно рассмотреть обычную MIDI-клавиатуру. Упрощенно, MIDI-клавиатура представляет собой укороченную клавиатуру рояля в корпусе с которой находится MIDI-интерфейс, позволяющий подключать ее к другим MIDI-устройствам, например, к MIDI-синтезатору, который установлен в звуковой карте компьютера. Используя специальное программное обеспечение (например, MIDI-секвенсор) можно включить MIDI-синтезатор в режим игры, например, на рояле, и нажимая на клавиши MIDI-клавиатуры слышать звуки рояля. Естественно, что роялем дело не ограничивается – в стандарте GM имеются 128 мелодических инструментов и 46 ударных. Кроме того, используя MIDI-секвенсор можно записывать исполняемые на MIDI-клавиатуре ноты в компьютер, для последующего редактирования и аранжировки, либо просто для элементарной распечатки нот.

Надо отметить, что поскольку MIDI-данные – это набор команд, то музыка, которая написана с помощью MIDI, также записывается с помощью команд синтезатора. Иными словами, MIDI-партитура – это последовательность команд: какую ноту играть, какой инструмент использовать, какова продолжительность и тональность ее звучания и так далее. Знакомые многим MIDI-файлы (.MID) есть нечто иное, как набор таких команд. Естественно, что поскольку имеется великое множество производителей MIDI-синтезаторов, то и звучать один и тот же файл может на разных синтезаторах по-разному (потому что в файле сами инструменты не хранятся, а есть лишь только указания синтезатору какими инструментами играть, в то время как разные синтезаторы могут звучать по-разному).

Вернемся к рассмотрению звукомузыкальных плат. Поскольку мы уже уточнили, что такое MIDI, нельзя обойти стороной характеристики встроенного аппаратного синтезатора звуковой карты. Современный синтезатор, чаще всего, основан на так называемой «волновой таблице» - WaveTable (вкратце, принцип работы такого синтезатора состоит в том, что звук в нем синтезируется из набора записанных звуков путем их динамического наложения и изменения параметров звучания), раньше же основным типом синтеза являлся FM (Frequency Modulation – синтез звука посредством генерирования простых синусоидальных колебаний и их смещения). Основными характеристиками WT-синтезатора являются: количество инструментов в ПЗУ и его объем, наличие ОЗУ и его максимальный объем, количество возможных эффектов обработки сигналов, а также возможность поканальной эффект-обработки (конечно, в случае наличия эффект-процессора), количество генераторов, определяющих максимальное число голосов в полифоническом (многоголосном) режиме и, может быть самое главное, стандарт, в соответствии с которым выполнен синтезатор (GM, GS или XG). Кстати, объем памяти синтезатора - не всегда величина фиксированная. Дело в том, что в последнее время синтезаторы перестали иметь свое ПЗУ, а пользуются основным ОЗУ компьютера: в этом случае все используемые синтезатором звуки хранятся в файле на диске и при необходимости считываются в ОЗУ.

7. Программное обеспечение

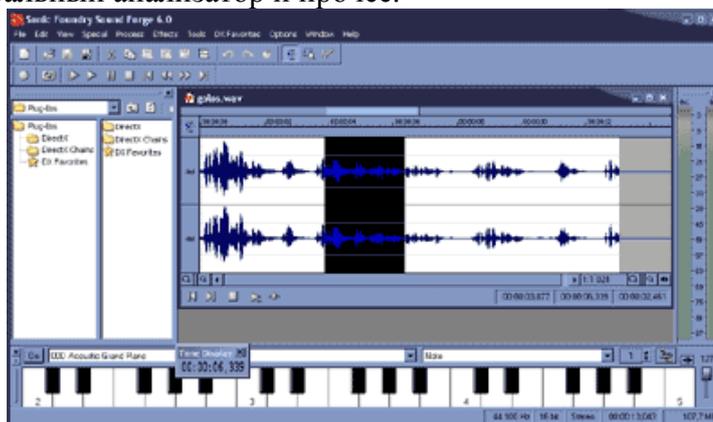
Тема программного обеспечения очень широка, поэтому здесь мы только вкратце обсудим основные представители программ для обработки звука.

Наиболее важный класс программ – редакторы цифрового аудио. Основные возможности таких программ это, как минимум, обеспечение возможности записи (оцифровки) аудио и сохранение на диск. Развитые представители такого рода программ позволяют намного больше: запись, многоканальное сведение аудио на нескольких виртуальных дорожках, обработка специальными эффектами (как встроенными, так и подключаемыми извне – об этом позже), очистка от шумов, имеют развитую навигацию и инструментарий в виде спектрографа и прочих виртуальных приборов, управление/управляемость внешними устройствами, преобразование аудио из формата в формат, генерация сигналов, запись на компакт диски и многое другое. Некоторые из таких программ: Cool Edit Pro (Syntrillium), Sound Forge (Sonic Foundry), Nuendo (Steinberg), Samplitude Producer (Magix), Wavelab (Steinberg).

Основные возможности редактора Cool Edit Pro 2.0 (см. Скриншот 1 - пример рабочего окна программы в многодорожечном режиме): редактирование и сведение аудио на 128 дорожках, 45 встроенных DSP-эффектов, включая инструменты для мастеринга, анализа и реставрации аудио, 32-битная обработка, поддержка аудио с параметрами 24 бит / 192 КГц, мощный инструментарий для работы с петлями (loops), поддержка DirectX, а также управление SMPTE/MTC, поддержка работы с видео и MIDI и прочее.



Основные возможности редактора Sound Forge 6.0a (см. Скриншот 2 - пример рабочего окна программы): мощные возможности не деструктивного редактирования, многозадачная фоновая обработка заданий, поддержка файлов с параметрами до 32 бит / 192 КГц, менеджер предустановок, поддержка файлов более 4 Гб, работа с видео, большой набор эффектов обработки, восстановление после зависаний, предпрослушивание примененных эффектов, спектральный анализатор и прочее.

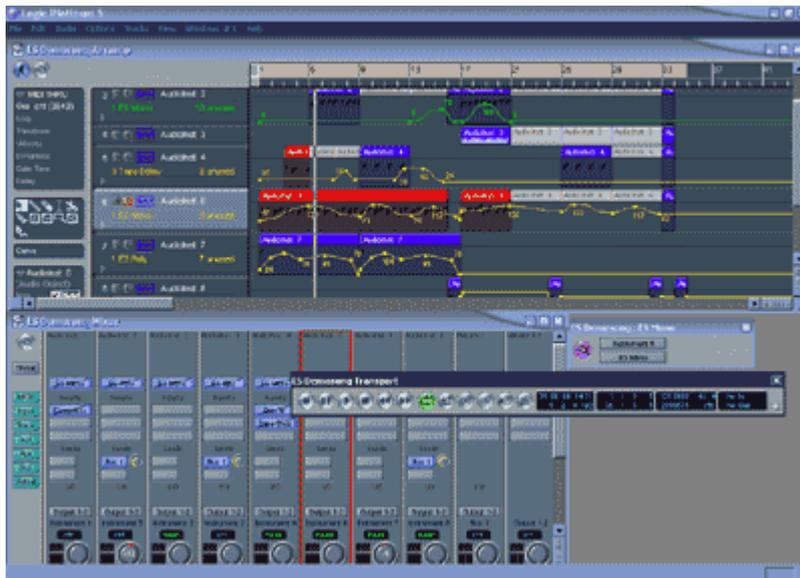


Не менее важная в функциональном смысле группа программ – секвенсоры (программы для написания музыки). Чаще всего, такие программы используют MIDI-синтезатор (аппаратный внешний или встроенный почти в любую звуковую карту, либо программный, организуемый специальным программным обеспечением). Такие программы предоставляют пользователю либо привычный нотный стан (как, например, программа Finale от CODA), либо более распространенный способ редактирования аудио на компьютере, так называемый, piano-roll (это более понятное представление музыки для людей, не знакомых с нотами; в таком представлении вертикально имеется ось с изображением клавиш пианино, а горизонтально откладывается время, таким образом, ставя на пересечении штрихи разной длины, добиваются звучания определенной ноты с определенной продолжительностью). Встречаются и программы, позволяющие просматривать и редактировать аудио в обоих представлениях. Развитые секвенсоры помимо редактирования аудио во многом могут дублировать возможности редакторов цифрового аудио – осуществлять запись на CD, совмещать MIDI-дорожки с цифровыми сигналами и осуществлять мастеринг. Яркие представители такого класса программ: Cubase (Steinberg), Logic Audio (Emagic), Cakewalk (Twelve Tone Systems) и уже упомянутый Finale.

Основные возможности редактора Cubase 5.1 (см. Скриншот 3 – пример рабочего окна программы в режиме просмотра MIDI дорожек): редактирование музыки в реальном времени используя графическое представление информации, высокое разрешение редактора (15360 пульсов на четверть), практически не лимитированное количество дорожек, 72 аудио канала, поддержка VST32, 4 эквалайзера на канал и другие поканальные эффекты, встроенные инструменты обработки с использованием аналогового моделирования (виртуальные инструменты, эффект процессоры, инструменты микширования и записи) и множество других возможностей.



Основные возможности редактора Logic Audio 5 (см. Скриншот 4 – пример рабочего окна программы): работа со звуком при точности в 32 бита, высокое временное разрешение событий, самоадаптируемый микшер аудио и MIDI, оптимизируемый интерфейс пользователя, синхронизация с видео, виртуально неограниченное число MIDI-дорожек, обработка звука в реальном времени, полная синхронизация с MTC, MMC, SMPTE, встроенные модули обработки и автоинструменты, поддержка большого количества аппаратного оборудования, а также множество других возможностей.

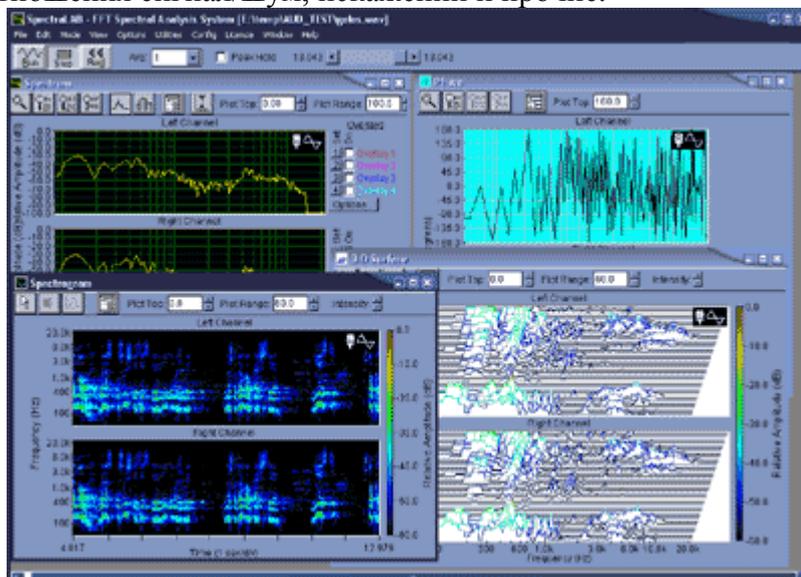


В наборе программ пользователя, занимающегося обработкой звука, имеется множество разных инструментов, так было раньше и так будет впредь – универсальных комбайнов для работы со звуком не бывает. Однако, не смотря на все разнообразие ПО, в программах часто используются схожие механизмы для обработки звука (например, процессоры эффектов и прочие). На каком-то этапе разработки аудио ПО, производители поняли, что удобнее сделать в своих программах возможность подключения внешних инструментов, чем каждый раз создавать заново инструменты для каждой отдельной программы. Так что многие программы, относящиеся к той или иной группе ПО, позволяют подключать так называемые «плагины» - внешние подключаемые модули, расширяющие возможности обработки звука. Это стало возможным в результате появления нескольких стандартов на интерфейс между программой и подключаемым модулем. На сегодняшний день существуют два основных стандарта на интерфейс: DX и VST. Существование стандартов позволяет подключать один и тот же плагин к совершенно разным программам, не заботясь о возникновении конфликтов и неполадок. Говоря о самих плагинах, надо сказать, что это просто огромное семейство программ. Обычно, один плагин является механизмом, реализующим какой-то конкретный эффект, например, реверберацию или низкочастотный фильтр. Из интересных плагинов можно вспомнить, например iZotope Vinyl, - он позволяет придать звучанию эффект виниловой пластинки (см. Скриншот 5 – пример рабочего окна плагина в среде Cool Edit Pro), Antares AutoTune позволяет в полуавтоматическом режиме корректировать звучание вокала, а Orange Vocoder являет собой замечательный вокодер (механизм для придания звучанию различных инструментов схожести со звучанием голоса человека).



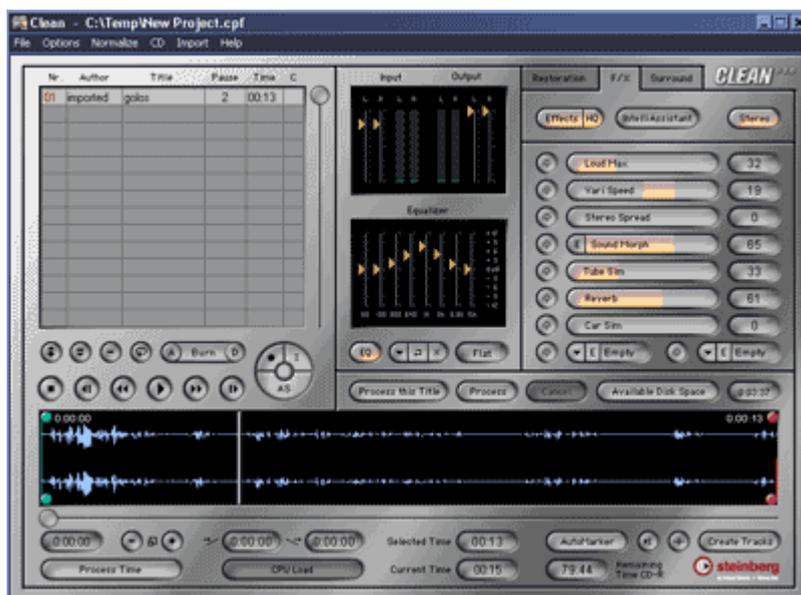
Обработка звука и написание музыки – это не только творческий процесс. Иногда нужен скрупулезный анализ данных, а также осуществление поиска огрехов их звучания. Кроме того, аудио материал, с который приходится иметь дело, не всегда желаемого качества. В этой связи нельзя не вспомнить о целом ряде программ-анализаторов аудио, специально предназначенных для осуществления измерительных анализов аудио данных. Такие программы помогают представить аудио данные удобнее, чем обычные редакторы, а также внимательно изучить их с помощью различных инструментов, таких как FFT-анализаторы (построители динамических и статических амплитудно-частотных характеристик), построители сонограмм, и прочих. Одна из наиболее известных и развитых программ подобного плана – программа SpectraLAB (Sound Technology Inc.), чуть более простые, но мощные – Analyzer2000 и Spectrogram.

Программа SpectraLAB – наиболее мощный продукт подобного рода, существующий на сегодня (см. Скриншот 6 – пример рабочего окна программы, на экране: спектральная картина в трех представлениях и фазовая картина). Возможности программы: 3 режима работы (пост режим, режим реального времени, режим записи), основной инструментарий – осциллограф, спектрометр (двухмерный, трехмерный, а также построитель сонограмм) и фазометр, возможность сравнения амплитудно-частотных характеристик нескольких сигналов, широкие возможности масштабирования, измерительные инструменты: нелинейных искажений, отношения сигнал/шум, искажений и прочие.



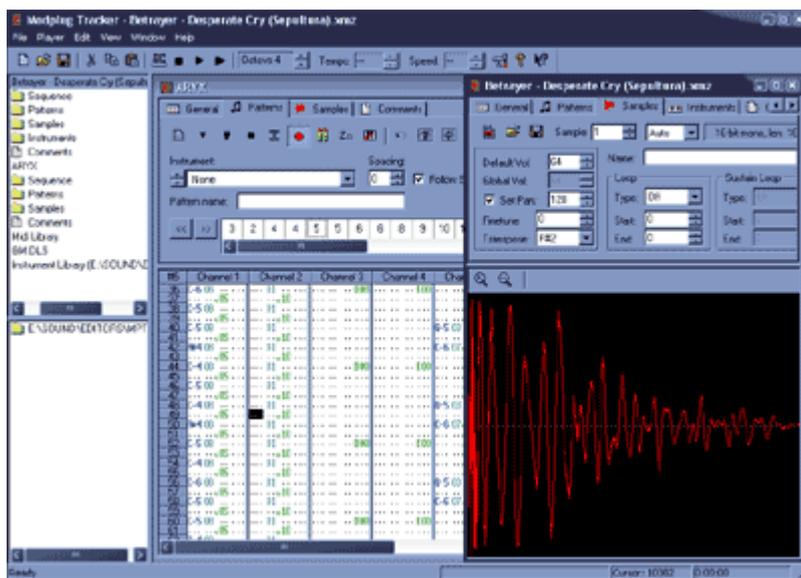
Специализированные реставраторы аудио играют также немаловажную роль в обработке звука. Такие программы позволяют восстановить утерянное качество звучания аудио материала, удалить нежелательные щелчки, шумы, треск, специфические помехи записей с аудио-кассет, и провести другую корректировку аудио. Программы подобного рода: Dart, Clean (от Steinberg Inc.), Audio Cleaning Lab. (от Magix Ent.), Wave Corrector.

Основные возможности реставратора Clean 3.0 (см. Скриншот 8 – рабочее окно программы): устранение всевозможных потрескиваний и шумов, режим автокоррекции, набор эффектов для обработки скорректированного звука, включая функцию «surround sound» с наглядным акустическим моделированием эффекта, запись CD с подготовленными данными, «интеллектуальная» система подсказок, поддержка внешних VST плагинов и другие возможности.



Трееры – это отдельная категория звуковых программ, предназначенных именно для создания музыки. Ранее мы рассмотрели два принципиально отличных способа хранения звуковых данных (музыки): первый - хранение звука в виде сжатого или несжатого потока аудио, второй - хранение музыки в виде MIDI-файлов (в виде набора команд MIDI-синтезатору). Структура и концепция построения трекерных файлов очень похожа на принцип хранения MIDI-информации. В трекерных модулях (файлы, созданные в трекерах, принято называть модулями), также, как и в MIDI-файлах, содержится партитура в соответствии с которой должны проигрываться инструменты. Кроме того, в них содержится информация о том, какие эффекты и в какой момент времени должны быть применены при проигрывании того или иного инструмента. Однако, принципиальное отличие трекерных модулей от MIDI-файлов заключается в том, что проигрываемые в этих модулях инструменты (или, точнее сказать, сэмплы) хранятся в самих модулях (то есть внутри файлов), а не в синтезаторе (как это происходит в случае с MIDI). Такой способ хранения музыки имеет массу преимуществ: размер файлов невелик по сравнению с непрерывной оцифрованной музыкой (поскольку записываются только использованные инструменты и партитура в виде команд), нет зависимости звучания от компьютера, на котором происходит воспроизведение (в MIDI, как мы говорили, есть зависимость звучания от используемого синтезатора), имеется большая свобода творчества, поскольку автор музыки не ограничен набором инструментов (как в MIDI), а может использовать в качестве инструмента любой оцифрованный звук. Основные программы-трекеры *Scream Tracker*, *Fast Tracker*, *Impulse Tracker*, *OctaMED*, *SoundStudio*, *MAD Tracker*, *ModPlug Tracker*.

Программа *ModPlug Tracker* является сегодня одним из тех трекеров, сумевших стать универсальной рабочей средой для множества типов трекерных модулей (см. Скриншот 7 – пример рабочего окна программы, на экране: содержание дорожек одного загруженного модуля и рабочее окно сэмплов другого модуля). Основные возможности: поддержка до 64 физических каналов аудио, поддержка почти всех существующих форматов трекерных модулей, импорт инструментов во множестве форматов, 32-битное внутреннее микширование, высококачественный ресэплирующий фильтр, поддержка MMX/3dNow!/SSE, автоматическое удаление потрескиваний, расширение басов, ревербератор, расширение стерео, 6-полосный графический эквалайзер и другие возможности.



Скриншот 7

Напоследок следует упомянуть о существовании огромного количества другого аудио ПО: проигрыватели аудио (наиболее выдающиеся: WinAMP, Sonique, Apollo, XMPlay, Cubic Player), подключаемые модули для проигрывателей (из «улучшателей» звучания аудио - DFX, Enhancer, iZotope Ozone), утилиты для копирования информации с аудио CD (ExactAudioCopy, CDex, AudioGrabber), перехватчики аудио потоков (Total Recorder, AudioTools), кодеры аудио (кодеры MP3: Lame encoder, Blade Encoder, Go-Go и другие; кодеры VQF: TwinVQ encoder, Yamaha SoundVQ, NTT TwinVQ; кодеры AAC: FAAC, PsyTel AAC, Quartex AAC), конвертеры аудио (для перевода аудио информации из одного формата в другой), генераторы речи и множество других специфических и общих утилит. Безусловно, все перечисленное – только малая толика из того, что может пригодиться при работе со звуком.

8. Перспективы и проблематика

Перспективы развития и использования цифрового аудио видятся авторам статьи очень широкими. Казалось бы, все, что можно было сделать в этой области, уже сделано. Однако это не так. Остается масса еще совсем незатронутых проблем.

Например, область распознавания речи еще очень не развита. Давно уже делались и делаются попытки создать программное обеспечение, способное качественно распознавать речь человека, однако все они пока не приводят к желаемому результату. А ведь долгожданный прорыв в этой области мог бы неимоверно упростить ввод информации в компьютер. Только представьте себе, что вместо набора текста его можно было бы просто надиктовывать, попивая кофе где-нибудь неподалеку от компьютера. Имеется множество программ якобы способных предоставить такую возможность, однако все они не универсальны и сбиваются при незначительном отклонении голоса читающего от заданного тона. Такая работа приносит не столько удобств, сколько огорчений. Еще куда более сложной задачей (вполне возможно, что и неразрешимой вовсе) является распознавание общих звуков, например, звучания скрипки в звуках оркестра или выделение партии рояля. Можно надеяться, что когда-нибудь такое станет возможным, ведь человеческий мозг легко справляется с такими задачами, однако сегодня говорить о хотя бы малейших сдвигах в этой области рано.

В области синтеза звука также есть пространство для изучения. Способов синтеза звука сегодня существует несколько, однако ни один из них не дает возможности синтезировать звук, который нельзя было бы отличить от настоящего. Если, скажем, звуки рояля или тромбона еще более-менее поддаются реализации, до правдоподобного звучания саксофона или электрогитары добиться еще так и не смогли – существует масса нюансов звучания, которые почти невозможно воссоздать искусственно.

Таким образом, можно смело сказать, что в области обработки, создания и синтеза звука и музыки еще очень далеко до того решающего слова, которое поставит точку на развитии этой отрасли человеческой деятельности.

Тема 2.2. Средства компьютерной аудиотехнологии

Распознавание речи — процесс преобразования речевого сигнала в цифровую информацию (например, текстовые данные). Обратной задачей является синтез речи.

Речевое общение является естественным и удобным для человека. Задача распознавания речи состоит в том, что бы убрать посредника в общении человека и компьютера. Управление машиной голосом в реальном времени, а также ввод информации посредством человеческой речи намного упростит жизнь современного человека. Научить машину понимать без посредника тот язык, на котором говорят между собой люди – задачи распознавания речи.

Впервые устройство, распознававшее речь, появилось в 1952, и было способно распознавать цифры, произнесённые человеком. В 1962 году на ярмарке компьютерных технологий в Нью-Йорке было представлено устройство IBM Shoebox. Коммерческие программы по распознаванию речи появились в начале девяностых годов. Обычно их используют люди, которые из-за травмы руки не в состоянии набирать большое количество текста. Эти программы (например, Dragon NaturallySpeaking, VoiceNavigator (англ.)) переводят голос пользователя в текст, таким образом, разгружая его руки. Надёжность перевода у таких программ не очень высока, но с годами она постепенно улучшается. Увеличение вычислительных мощностей мобильных устройств позволило и для них создать программы с функцией распознавания речи. Среди таких программ стоит отметить приложение Microsoft Voice Command, которое позволяет работать со многими приложениями при помощи голоса. Например, можно включить воспроизведение музыки в плеере или создать новый документ. Все большую популярность применение распознавания речи находит в различных сферах бизнеса, например, врач в поликлинике может проговаривать диагнозы, которые тут же будут внесены в электронную карточку. Или другой пример. Наверняка каждый хоть раз в жизни мечтал с помощью голоса выключить свет или открыть окно. В последнее время в телефонных интерактивных приложениях все чаще стали использоваться системы автоматического распознавания и синтеза речи. В этом случае общение с голосовым порталом становится более естественным, так как выбор в нём может быть осуществлен не только с помощью тонового набора, но и с помощью голосовых команд. При этом системы распознавания являются независимыми от дикторов, то есть распознают голос любого человека. Следующим шагом технологий распознавания речи можно считать развитие так называемых интерфейсов безмолвного доступа (silent speech interfaces, SSI). Эти системы обработки речи базируются на получении и обработке речевых сигналов на ранней стадии артикулирования. Данный этап развития распознавания речи вызван двумя существенными недостатками современных систем распознавания: чрезмерная чувствительность к шумам, а также необходимость четкой и ясной речи при обращении к системе распознавания. Подход, основанный на SSI, заключается в том, чтобы использовать новые сенсоры, не подверженные влиянию шумов в качестве дополнения к обработанным акустическим сигналам.

Системы распознавания речи можно классифицировать в зависимости от:

- назначения (системы диктовки, командные системы);
- типа речи (слитная или раздельная речь);
- размера словаря (ограниченный набор слов, словарь большого размера);
- диктора (дикторозависимые и дикторонезависимые системы);
- механизма функционирования (простейшие (корреляционные) детекторы, экспертные системы с различным способом формирования и обработки базы знаний, вероятностно-сетевые модели принятия решения, в том числе нейронные сети);
- используемого алгоритма (нейронные сети, скрытые Марковские модели, динамическое программирование);

типа структурной единицы (фразы, слова, фонемы, дифоны, аллофоны);
 принципа выделения структурных единиц (распознавание по шаблону, выделение лексических элементов).

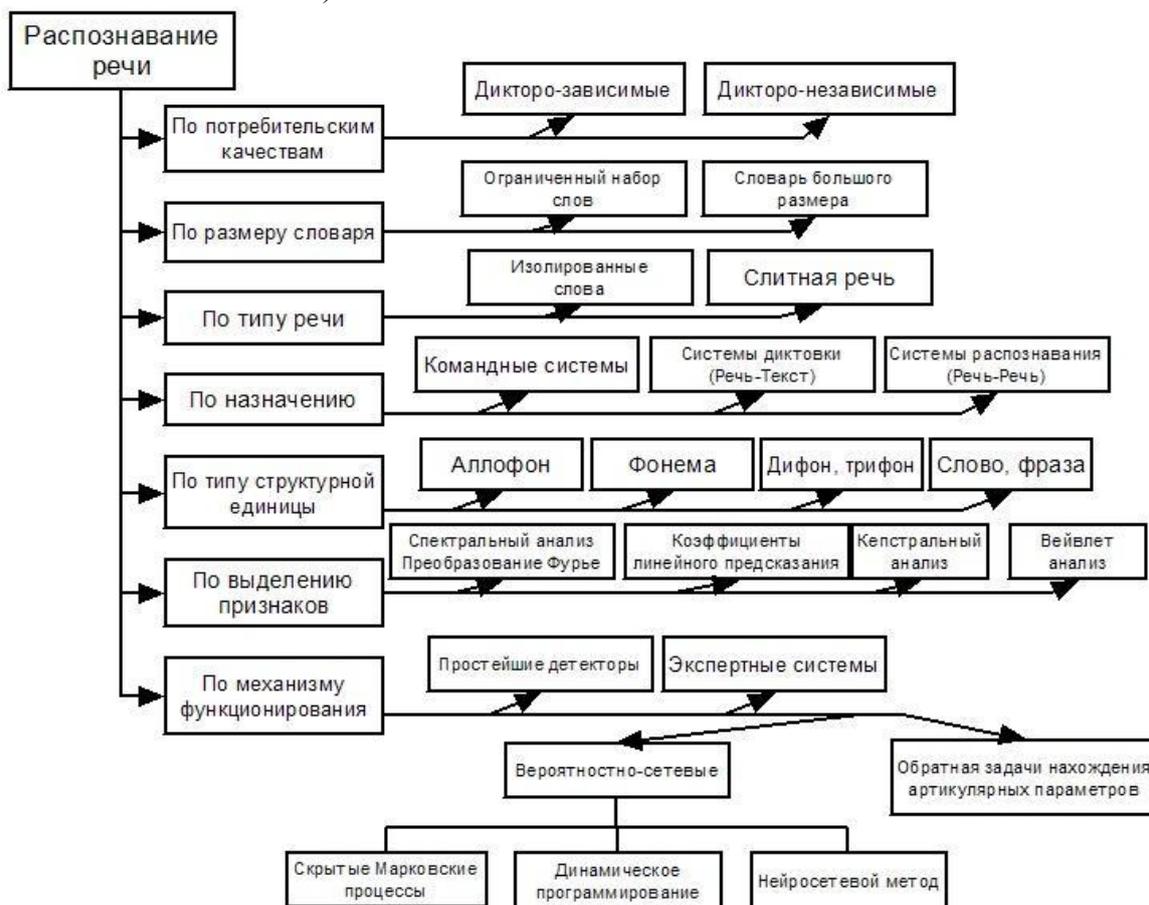


Рисунок Различные классификации систем распознавания речи.

Для систем автоматического распознавания речи, помехозащищённость обеспечивается, прежде всего, использованием двух механизмов:

Использование нескольких, параллельно работающих, способов выделения одних и тех же элементов речевого сигнала на базе анализа акустического сигнала;

Параллельное независимое использование сегментного (фонемного) и целостного восприятия слов в потоке речи.

Для систем автоматического распознавания речи, помехозащищённость обеспечивается, прежде всего, использованием двух механизмов:

Использование нескольких, параллельно работающих, способов выделения одних и тех же элементов речевого сигнала на базе анализа акустического сигнала;

Параллельное независимое использование сегментного (фонемного) и целостного восприятия слов в потоке речи.

В настоящее время распознавание речи сводится к решению трех типов задач:

- распознавание отдельно произносимых слов (используется для речевого управления вычислительной машиной);

- распознавание слитной речи (имеет целью преобразования в текст естественной речи человека);

- идентификация по образцу речи (используется для целей обеспечения безопасности).

Она состоит из трех стадий: регистрации, тестирования и допуска.

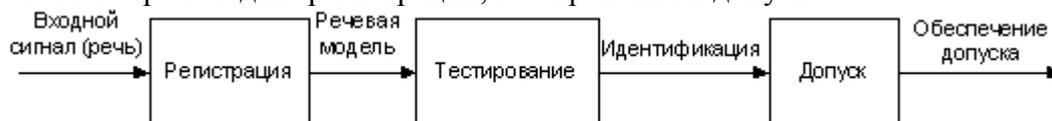


Рисунок Схема идентификации по образцу речи.

В процессе регистрации пользователя запоминаются особенности его голоса и формируется так называемая речевая модель. При тестировании выполняется сравнение предложенного образца речи с запомненной речевой моделью пользователя, а также с моделью "самозванца", составленной на базе голосов множества других людей. Если результат сравнения окажется положительным для первого случая и отрицательным для второго, считается, что тестирование прошло успешно.

Основные цели программных пакетов идентификации личности по голосу заключаются в следующем:

использование компьютера для автоматической сегментации речевого сигнала и автоматического измерения параметров;

повышение надежности идентификации за счет комплексного анализа речевого сигнала;

уменьшение субъективности эксперта и времени проведения экспертизы за счет автоматической обработки информации.

В большинстве существующих механизмов можно выделить четыре основных модуля:

Модуль сбора данных - включает получение входного сигнала и его предварительную обработку, которая может включать автоматическую регулировку усиления, подавление эха, обнаружение присутствия/отсутствия речи и обнаружение интонационного конца фразы. Этот модуль включает также выделение отрезка речи из входящего речевого сигнала.

Экстрактор - выполняет частотный анализ сигнала. Акустическо-фонетический поток данных разбивается на короткие кадры, или векторы, продолжительностью около 10 мс. Как правило, для каждого кадра определяется ряд параметров, используя быстрое преобразование Фурье. Кроме того, многие системы используют вместо или вместе с этими характеристиками другие, например, спектральные характеристики, а также первую и вторую производную от спектральных характеристик.

Компаратор - осуществляет акустическое сравнение: каждый кадр, или вектор, сравнивается с имеющимися акустическо-фонетическими образцами, хранящимися в специальной базе данных. При этом сравниваться могут как отдельные фонемы, так и слова, и даже фразы. При небольшом количестве слов, используемых диктором, более высокую надежность и скорость можно ожидать от распознавания целых слов, но при увеличении словаря скорость резко падает, и оптимальным становится распознавание отдельных фонем.

Интерпретатор решает задачу динамического программирования с целью найти наилучшее разбиение полученного от компаратора "алфавитного" потока на слова и фразы. В зависимости от объема используемого словаря и действующих синтаксических правил, применяются различные стратегии поиска и отсева.

В данном блоке из распознанных фонем формируются слова, а из слов фразы. При этом также часто используется балльная (или вероятностная) система сравнения результатов. Например, может использоваться алгоритм лучевого поиска по Витерби для определения наиболее вероятного предложения.

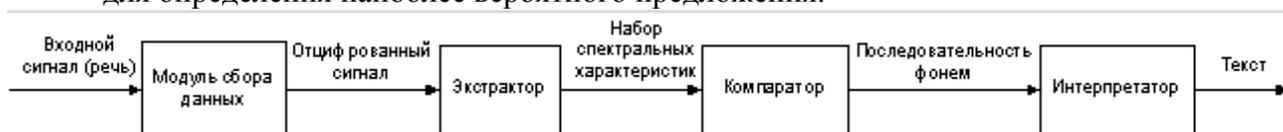


Рис.4. Схема механизма распознавания речи.

Классификация методов распознавания речи на основе сравнения с эталоном:

Динамическое программирование — временные динамические алгоритмы (Dynamic Time Warping).

Контекстно-зависимая классификация - при её реализации из потока речи выделяются отдельные лексические элементы — фонемы и аллофоны, которые затем объединяются в слоги и морфемы.

Методы дискриминантного анализа, основанные на Байесовской дискриминации (Bayesian discrimination);

Скрытые Марковские модели (Hidden Markov Model);

Нейронные сети (Neural networks).

Существует несколько алгоритмов определения начала и конца речи.

В одном из них определяется некоторый пороговый уровень сигнала. Начальная точка речи в этом случае соответствует моменту, когда входящий сигнал начинает превышать пороговый уровень, а конечная точка – моменту, где амплитуда входящего сигнала меньше пороговой. Основной недостаток этого метода заключается в невозможности точного определения речевого отрезка в случае сильного шума, или, наоборот, тихой речи.

Другой метод использует нормализацию амплитуды входного сигнала в соответствии с минимальной амплитудой. Полученные нормализованные значения сравниваются с пороговым значением.

Системы автоматической обработки устной речи находят практическое применение в информационно-справочных службах, где можно получать информацию из базы данных в режиме диалога (например, в медицине или на транспорте). Кроме того, такие системы необходимы и для организации приема и озвучивания сообщений (например, получение электронной почты по телефону), а также для перевода звучащей речи в привычный текст в электронной форме. Компьютеры могут оказывать помощь и при обучении иностранному языку с помощью автоматических фонетических тренажеров.

История практического применения систем автоматической обработки звучащей речи началась еще в XVIII в., когда появились первые механические синтезаторы речи. Их создатели ставили целью воспроизвести процессы произнесения звуков с помощью механического устройства, имитируя строение голосового аппарата человека.

В начале XX века механические устройства сменились электрическими вокодерами. Первое устройство для распознавания речи появилось в 1952 г., оно могло распознавать произнесенные человеком цифры. В 1964 г. на ярмарке компьютерных технологий в Нью-Йорке было представлено устройство IBM Shoebox.

Коммерческие программы по распознаванию речи появились в начале 90-х годов. Обычно их используют люди, которые из-за травмы руки не в состоянии набирать большое количество текста. Эти программы (н-р, Dragon NaturallySpeaking, VoiceNavigator) переводят голос пользователя в текст, таким образом, разгружая его руки. Надёжность перевода у таких программ не очень высока, но с годами она постепенно улучшается.

Увеличение вычислительных мощностей мобильных устройств позволило и для них создать программы с функцией распознавания речи. Среди таких программ стоит отметить приложение Microsoft Voice Command, которое позволяет работать со многими приложениями при помощи голоса. Например, можно включить воспроизведение музыки в плеере или создать новый документ.

Прогресс, однако, не стоит на месте и в последнее время в телефонных интерактивных приложениях все чаще стали использоваться системы автоматического распознавания и синтеза речи. В этом случае общение с голосовым порталом становится более естественным, так как выбор в нем может быть осуществлен не только с помощью тонового набора, но и с помощью голосовых команд. При этом системы распознавания являются независимыми от дикторов, то есть распознают голос любого человека.

Следующим шагом технологий распознавания речи можно считать развитие так называемых Silent Speech Interfaces (SSI) (Интерфейсов Безмолвного Доступа). Эти системы обработки речи базируются на получении и обработке речевых сигналов на ранней стадии артикулирования. Данный этап развития распознавания речи вызван двумя существенными недостатками современных систем распознавания: чрезмерная чувствительность к шумам, а также необходимость четкой и ясной речи при обращении к системе распознавания. Подход, основанный на SSI, заключается в том, чтобы использовать новые сенсоры, не подверженные влиянию шумов в качестве дополнения к обработанным акустическим сигналам.

На сегодняшний день существует два типа систем распознавания речи – 1) работающие по принципу «клиент-сервер» (client-server), 2) «на клиенте» (client-based). При использовании клиент-серверной технологии речевая команда вводится на устройстве пользователя и через Интернет передается на удаленный сервер, где обрабатывается и возвращается на устройство в виде команды (Google Voice, Vlingo); ввиду большого количества пользователей сервера система распознавания получает большую базу для обучения.

Первый вариант работает на иных математических алгоритмах и встречается редко (Speereo Software) – команда вводится на устройстве пользователя и обрабатывается в нем же. Плюс обработки «на клиенте» в мобильности, независимости от наличия связи и работы удаленного оборудования. Так, система, работающая «на клиенте» кажется надежнее, но иногда ограничивается мощностью устройства на стороне пользователя.

§ 4.3. Структура программ распознавания и синтеза звучащей речи

Современные синтезаторы речи включают два блока: блок лингвистической обработки текста, с помощью которого строится полная фонетическая транскрипция синтезируемого текста, а также блок акустического синтеза, который генерирует речевой сигнал.

Блок лингвистической обработки текста имеет достаточно сложную структуру, поскольку создание транскрипции включает несколько этапов: определение языка входного текста, устранение возможных орфографических ошибок, проведение морфологического анализа словоформ для постановки ударения. Самая трудная задача этапа лингвистической подготовки текста – формирование интонации и просодических характеристик фразы. Во многих случаях для этого необходим значительно более сложный семантический и синтаксический анализ фразы. Последний этап работы блока лингвистической подготовки текста – создание фонетической транскрипции. На этом этапе применяются стандартные правила чтения, при этом сложность и трудоемкость этого этапа определяется соотношением между орфографией и произношением каждого конкретного языка.

После создания фонетической транскрипции начинает работу второй блок синтезатора блок акустического синтеза. Его задача – перевод транскрипции в цифровой сигнал, который, в свою очередь, преобразуется в звуковые колебания при помощи обычного цифро-аналогового преобразователя.

Dragon Naturally Speaking – это мировой лидер в программном обеспечении по распознаванию человеческой речи. Программа дает большие возможности при использовании компьютера. Пользователь может диктовать тексты в микрофон, и программа будет писать их сама, например, в текстовом процессоре.

Программные решения синтеза русской и английской речи, а также программные комплексы распознавания английской речи предлагаются следующими компаниями:

Sakrament TTS (Text-to-Speech) Engine – система нового поколения, осуществляющая качественный речевой синтез. Она может использоваться как отдельное приложение для озвучивания электронных текстов, в качестве речевого движка для других приложений, а также для интеграции с различными информационными системами.

Sakrament ASR Engine – разработка компании «Сакрамент», рассчитана на применение в различных аппаратных системах и программных приложениях, использующих технологии распознавания речи, таких как: IVR-системы, мобильные электронные устройства, бытовая техника и т.д. Sakrament ASR Engine может быть легко перенесена на любую существующую программную или аппаратную платформу.

РАЗДЕЛ 3. КОМПЬЮТЕРНЫЕ ВИДЕОТЕХНОЛОГИИ

Тема 3.1. Ввод видеоданных в персональный компьютер

Цифровая обработка видеoinформации является одним из важнейших направлений в информационных технологиях, служащая для реализации функций искусственного интеллекта, связанных с обработкой статических изображений и видеопотоков. Видеоаналитика – раздел цифровой обработки видеoinформации, который позволяет получать из видеопотока

информацию, относящуюся к изображению в целом или к отдельным его элементам, а также улучшать визуальное восприятие видеопотока.

Аналоговый и цифровой видеосигналы

Аналоговый видеосигнал стандартного разрешения. Под видеосигналом понимается электрический сигнал специальной формы, посредством которого передается телевизионное изображение. Видеосигнал, принятый для использования в настоящее время в РФ описывается ГОСТ 7845-92. Стандарт регламентирует параметры и характеристики используемого телевизионного изображения и электрического сигнала, используемого для передачи от источника телевизионного изображения к приемнику. Это позволяет добиться совместимости телевизионных устройств по видеосигналу. Источником видеосигнала является формирователь видеосигнала. В качестве последнего могут выступать телевизионные камеры любого типа, устройства видеозаписи (видеомагнитофоны, DVD-проигрыватели, компьютеры с ТВ видеовыходами и пр.). Приемником видеосигнала являются устройства отображения видеoinформации, в частности телевизоры и видеомониторы, работающие на любом принципе (ЭЛТ, ЖК, плазменная панель), а также устройства видеозаписи (видеомагнитофоны, записывающие DVD-проигрыватели, видеорегистраторы систем видеонаблюдения). Мгновенное значение освещенности на фоточувствительной поверхности телевизионного фотоприемника преобразуется в мгновенное значение напряжения на выходе этого фотоприемника. Таким образом, в видеосигнале напряжение пропорционально яркости в данной точке изображения. Эта часть видеосигнала носит название сигнала яркости и используется для передачи черно-белого видеосигнала или сигнала яркости в цветном видеосигнале. Диапазон передаваемых значений яркости определяется уровнем черного и уровнем белого сигналов в видеосигнале. Уровень черного является минимальным сигналом яркости и соответствует уровню синхроимпульсов. Он же одновременно является и опорным сигналом. Уровень белого соответствует максимальному уровню передаваемой яркости (Рис. 1). Кроме сигнала яркости в видеосигнале присутствуют служебные составляющие, обеспечивающие синхронизацию сигнала между источником и приемником. Существуют два типа сигналов синхронизации – строчные и кадровые. Сигналы кадровой синхронизации обеспечивают передачу информации о времени начала каждого поля телевизионного изображения в видеосигнале, а также информацию о типе этого поля (четное или нечетное). В соответствии с период импульсов кадровой синхронизации составляет 20 мс.

Форматы представления видеосигнала

Низкочастотный телевизионный видеосигнал является композитным, т.е. представляет собой результат сложения яркостного сигнала Y , двух цветowych поднесущих, модулированных сигналами цветности U и V , а также синхроимпульсов, причем частоты цветоразностных сигналов лежат в 79 пределах полосы спектра яркостного сигнала. В бытовых устройствах ограничиваются более простыми полосовыми фильтрами, заметно снижающими четкость изображений. Так в видеомагнитофонах и камерах классов VHS (Video Home System) и Video-8 используются только композитные видеосигналы, при этом разрешение ограничено 240 телевизионными строками. Следующим шагом к повышению качества является переход к компонентному сигналу YUV . Он используется в профессиональной аппаратуре класса Betacam и обеспечивает разрешение до 500 строк. Последним шагом в этой череде является RGB-представление, при котором отсутствуют какое-либо кодирование и модуляция, и осуществляется наиболее простая и точная передача сигнала. Подобное представление реально используется только в высокоточной научной измерительной аппаратуре. За последние несколько лет появилось большое число различных цифровых форматов представления видеосигнала. Аппаратура, работающая в этих форматах, выпускается рядом фирм – законодателей мод в видеотехнике, такими как Sony, Panasonic, JVC и т.д.

Основные методы компрессии цифрового видеосигнала

Несмотря на огромный поток данных, составляющий цифровой видеосигнал, существуют предпосылки для эффективной его компрессии. Дело в том, что видеосигнал обладает очень большой избыточностью, которую следует использовать при компрессии. Некото-

рые методы сокращения избыточности используются еще в аналоговом видеосигнале. К таким методам можно отнести: - значительное снижение полосы частот сигналов цветности; - использование чересстрочной развертки; - удаление части спектра сигнала яркости с помощью фильтра пробки и включение в освободившуюся полосу сигнала цветности. Эти методы позволили сократить полосу частот аналогового видеосигнала и использовать один и тот же частотный диапазон как для черно-белого, так и для цветного видеосигнала. В цифровом виде для компрессии используются следующие предпосылки: -

- когерентность областей изображения, т.е. малое изменение цвета изображения в соседних пикселах;

- избыточность в цветовых плоскостях, т.е. использование большей важности яркости изображения по сравнению с цветностью (фактически это следствие уменьшение полосы частот сигнала цветности);

- подобие между кадрами, т.е. соседние кадры изменяются незначительно.

Первые два пункта используются не только для компрессии видеосигнала, но цифровых изображений. Общей особенностью как для изображений, так и для видео является тот факт, что человеческое зрение при анализе изображения оперирует контурами, общим переходом цветов и сравнительно малочувствительно к малым измерениям в изображении. Следовательно, возможно создание эффективных алгоритмов компрессии, в которых декомпрессированное изображение или видео не будет совпадать с оригиналом, но человек этого не заметит. Такая особенность человеческого зрения позволила создать специальные алгоритмы сжатия, ориентированные только на изображение и видео. Изображение обладает избыточностью в двух измерениях. То есть, обычно соседние точки как по горизонтали, так и по вертикали в изображении близки по цвету. Кроме того, можно использовать подобие между цветовыми плоскостями R, G, B или Y, U, V. Таким образом, при создании алгоритмов компрессии изображений и видео используются особенности структуры изображения и особенности человеческого зрения.

На данный момент известны три семейства специальных алгоритмов для сжатия изображений и видео и метод межкадрового сжатия в видео, которые невозможно применить к архивации еще каких-либо видов данных. Кроме того, методы сжатия можно разделить на два основных типа: - без потери информации; - с потерей информации. К первому типу можно отнести такие методы, которые после декомпрессии позволяют получить исходную информацию без искажений. Такие методы используются во всех известных архиваторах данных, а также для сжатия изображений и видео, потери в которых крайне нежелательны, например, в медицинских приложениях. Однако, такие методы обычно обладают небольшим коэффициентом сжатия. Второй тип включает в себя методы компрессии, при которых декомпрессированные данные отличаются от исходных. Но эти отличия возможно контролировать и удерживать на допустимом уровне. Основные алгоритмы компрессии изображений и видео используют именно эти методы. Поскольку эти алгоритмы обычно состоят из нескольких методов, то в их число могут входить и методы сжатия без потери информации.

Цифровая обработка видеoinформации является одним из важнейших направлений в информационных технологиях, служащая для реализации функций искусственного интеллекта, связанных с обработкой статических изображений и видеопотоков.

Видеосигнал, принятый для использования в настоящее время в Российской Федерации, описывается по ГОСТ 7845-92. Стандарт регламентирует параметры и характеристики используемого телевизионного изображения и электрического сигнала, используемого для передачи от источника телевизионного изображения к приемнику. Это позволяет добиться совместимости телевизионных устройств по видеосигналу.

В случае использования цветного изображения в видеосигнал дополнительно включаются сигналы цветности и цветовой синхронизации. В настоящее время используются три стандарта цветного видеосигнала - SECAM, PAL и NTSC.

Цифровой видеосигнал получается из аналогового видеосигнала путем преобразования его в цифровую форму. Это преобразование включает в себя три операции:

1. Дискретизацию во времени, т.е. преобразование непрерывного во времени видеосигнала в последовательность отсчетов, обновляющихся через определенные промежутки времени.

2. Квантование по уровню - замена непрерывной шкалы амплитуды видеосигнала на выбранный набор уровней квантования с округлением значения каждого отсчета до ближайшего уровня квантования.

3. Кодирование отсчетов, в результате которого значение отсчета представляется в виде числа, соответствующего номеру полученного уровня квантования.

При обработке и передаче видеосигнала в цифровом виде требуется его перевод в цифровую форму.

Сначала цифровой видеосигнал использовался для передачи телевизионного изображения стандартного разрешения. С таким разрешением работает и аналоговый видеосигнал. Поэтому цифровой видеосигнал на первых порах рассматривался как цифровой эквивалент стандартного аналогового видеосигнала. Фактически необходимо было иметь возможность производить аналого-цифровое и цифроаналоговое преобразование стандартного видеосигнала таким образом, чтобы при преобразовании не было потерь качества телевизионного изображения. С этой целью был разработан стандарт цифрового видеосигнала стандартного разрешения ITU-R BT.601 (он же CCIR-601), а немного позже вышла его обновленная версия ITU-R BT.656 (он же CCIR-656). Основными параметрами цифрового сигнала являются частота дискретизации и число уровней квантования по амплитуде. Последний параметр тесно связан с разрядностью используемых АЦП и ЦАП.

Несмотря на огромный поток данных, составляющий цифровой видеосигнал, существуют предпосылки для эффективной его компрессии. Дело в том, что видеосигнал обладает очень большой избыточностью, которую следует использовать при компрессии. Некоторые методы сокращения избыточности используются еще в аналоговом видеосигнале. К таким методам можно отнести:

- ◆ значительное снижение полосы частот сигналов цветности;
- ◆ использование чересстрочной развертки;
- ◆ удаление части спектра сигнала яркости с помощью фильтра пробки и включение в освободившуюся полосу сигнала цветности.

Эти методы позволили сократить полосу частот аналогового видеосигнала и использовать один и тот же частотный диапазон как для черно-белого, так и для цветного видеосигнала.

В цифровом виде для компрессии используются следующие предпосылки [3]:

- ◆ когерентность областей изображения, т.е. малое изменение цвета изображения в соседних пикселах;
- ◆ избыточность в цветовых плоскостях, т.е. используется большая важность яркости изображения по сравнению с цветностью (фактически это следствие уменьшения полосы частот сигнала цветности);
- ◆ подобие между кадрами, т.е. соседние кадры изменяются незначительно.

Кроме того, методы сжатия можно разделить на два основных типа:

- 1) без потери информации;
- 2) с потерей информации.

К первому типу можно отнести такие методы, которые после декомпрессии позволяют получить исходную информацию без искажений. Такие методы используются во всех известных архиваторах данных, а также для сжатия изображений и видео, потери в которых крайне нежелательны, например, в медицинских приложениях. Однако такие методы обычно обладают небольшим коэффициентом сжатия.

Второй тип включает в себя методы компрессии, при которых декомпрессированные данные отличаются от исходных. Но эти отличия возможно контролировать и удерживать на допустимом уровне. Основные алгоритмы компрессии изображений и видео используют

именно эти методы. Поскольку эти алгоритмы обычно состоят из нескольких методов, то в их число могут входить и методы сжатия без потери информации.

Среди основных методов, реализующих пространственно-частотное преобразование, чаще всего используется дискретное косинусное преобразование (ДКП) или дискретное вейвлет-преобразование (ДВП). Причем первое ДВП было предложено венгерским математиком Альфредом Хааром. Для входного сигнала, представленного массивом 2П чисел, вейвлет преобразование Хаара просто группирует элементы по 2 и образует от них суммы и разности. Группировка сумм проводится рекурсивно для образования следующего уровня разложения. В итоге получается 2п-1 разность и 1 общая сумма.

Основные задачи и методы видеоаналитики. Видеоаналитикой называют технологию компьютерного анализа видеоданных с целью получения систематизированной информации об объекте наблюдения без участия оператора. Как правило, видеоаналитика представляет собой программные алгоритмы, которые можно запускать как на компьютерах, так и встраивать в камеры.

Как известно, видеоинформационные системы подразделяются на два больших класса:

1) визуальные системы, приемником информации с которых является человек со своей зрительной системой;

2) системы технического зрения, информация с которых обрабатывается автоматической системой.

В первом случае основная задача видеоаналитики - помочь человеку воспринимать поступающую видеоинформацию. Для этого может потребоваться решение следующих частных задач, связанных с улучшением зрительного восприятия:

- ◆ повышение качества изображения (снижение видимого шума на изображении, повышение контраста изображения, улучшение четкости);
- ◆ улучшение цветопередачи или введение псевдоцвета;
- ◆ акцентирование внимания на отдельных важных деталях изображения;
- ◆ интерполяция (увеличение количества) пикселей, строк и кадровой частоты;
- ◆ перевод видеопотока с чересстрочной разверткой в построчную (прогрессивную) и наоборот;

◆ адаптация работы камеры к плохим условиям работы (низкая освещенность, наличие тумана и других атмосферных воздействий).

Во втором случае с помощью видеоаналитики решаются задачи автоматического выделения и обработки информации из видеопотока. К таким задачам можно отнести следующие:

- ◆ детектор движения в видеопотоке;
- ◆ детектор оставленных предметов;
- ◆ распознавание букв, цифр и других символов и состоящего из них текста;
- ◆ выделение человеческих фигур;
- ◆ выделение лиц;
- ◆ распознавание лиц;
- ◆ наблюдение за животными;
- ◆ измерение геометрических параметров объектов;
- ◆ совместная обработка нескольких видеопотоков (комплексирование видеоинформации).

Одной из наиболее сложных и актуальных задач обработки видеоизображения является проблема выделения и распознавания движущихся объектов при наличии различного рода помех и создание на этой основе системы мониторинга. Главная задача таких систем - информировать человека о ситуации, сложившейся в поле зрения камеры, и по возможности предпринять какие-либо заранее предусмотренные и программно заложенные действия [6].

На первых порах задача сводилась к простому детектированию движущихся объектов. Уже в таком виде это важно для многих систем безопасности, когда в охраняемой зоне ис-

ключено какое бы то ни было движение. Более совершенные системы подразумевают наличие в своем составе интеллектуальных видеодетекторов движения, способных отличить движущегося человека от собаки, машины или дерева, раскачивающегося на ветру. Привлекая внимание к определенному монитору и автоматически регистрируя произошедшее событие, такие системы значительно повышают уровень безопасности охраняемого объекта. Кроме того, видеодетекторы движения часто используются для интеллектуальной компрессии, что позволяет значительно экономить дисковое пространство при архивировании видеоизображений.

Условно обработку видеоизображений в таких системах можно разделить на следующие этапы:

- ◆ выделение переднего плана;
- ◆ выделение и классификация движущихся объектов;
- ◆ отслеживание траектории движения найденных объектов;
- ◆ распознавание и описание действий объектов, представляющих интерес.

Процесс выделения переднего плана заключается в отделении движущихся фрагментов изображения от неподвижных, которые называют фоновыми или принадлежащими заднему плану. От того, насколько аккуратно и корректно решена эта задача, зависят все последующие этапы обработки информации, а также требуемые вычислительные ресурсы. Сложность этой задачи обуславливается большим количеством разнообразных факторов, таких как собственные шумы камеры, внезапное изменение освещенности сцены, падающие тени, движение ветвей деревьев на ветру и др.

На втором этапе выделения и классификации сначала производится сегментация изображения переднего плана, т.е. находятся компактные области, движущиеся с одинаковой скоростью, которые считаются движущимися объектами. Далее они соотносятся с заранее определенными классами.

Методы построения переднего плана. Для построения переднего плана наиболее часто используются методы, основанные на вычитании фона, вероятностные методы, методы временной разности и оптического потока.

Методы вычитания фона являются самыми простыми и наиболее часто применяемыми для детектирования движущихся объектов. Суть их заключается в нахождении попиксельной разности между текущим кадром и некой моделью фона. В принципе, такая модель должна представлять собой сцену без движущихся объектов. При этом необходимо ее регулярное обновление, для того чтобы учитывать изменение условий освещенности и настроек камеры, таких как поворот, наклон и изменение фокусного расстояния. Главным недостатком методов вычитания фона является возможная классификация фоновых пикселей как переднеплановых. Это может происходить, например, для листьев деревьев, колышущихся на ветру, падающих снега и дождя, теней, отбрасываемых движущимися объектами и др. Кроме того, методам данного класса присуща латентность в обновлении модели фона: должно пройти некоторое время, прежде чем в модели будут учтены изменения, связанные с началом движения или остановкой объекта. Наконец, методы вычитания фона в своей простейшей реализации предъявляют достаточно высокие требования к ресурсам вычислительной системы.

В настоящее время развиваются методы построения заднего плана, основанные на применении искусственных нейронных сетей. Эти методы используют свойство нейронной сети адаптироваться к входным данным за счет введения настраиваемых обратных связей. Каждый пиксел фона управляется своей нейронной сетью, в результате чего через некоторое время, требуемое для настройки (обучения) нейронной сети, формируется модель фона, способная заданным образом подстраиваться к изменениям входного изображения.

В качестве итога можно сказать, что был произведен ознакомительный анализ существующих методов цифровой обработки видеосигналов. Даны базовые определения понятий, связанных с цифровой обработкой речевых сигналов, описаны основные методы обработки

видеосигналов и представлены примеры использования данных методов в зависимости от постановки конкретной

Тема 3.2. Особенности обработки цифровой видеoinформации

Видеомонтаж – это процесс "сборки" фильма из отдельных кадров. Основные задачи видеомонтажа – удаление ненужных участков сюжета, состыковка отдельных фрагментов видеоматериала, создание переходов между ними, добавление спецэффектов и поясняющих титров. Различают: линейный, нелинейный и гибридный, внутрикадровый и межкадровый, последовательный и параллельный монтаж.

Технология линейного монтажа основана на перезаписи видеоматериала с нескольких видеоисточников на записывающее устройство с удалением ненужных и "вклеиванием" нужных видеосцен и спецэффектов. Это технология прошлого.

В настоящее время используется технология нелинейного монтажа видеоматериала, осуществляемая с помощью цифровых компьютерных технологий.

Гибридный монтаж применяется в особых случаях и использует технологии комбинации линейного и нелинейного.

Внутрикадровый монтаж – это разные приемы съемки в течение одного кадра, выбираемые оператором на камере: "переход фокуса", "панорама", "наезд", "отъезд", "движение камеры".

Межкадровый монтаж заключается во включении блока специальных кадров, обеспечивающих плавный переход одного события к другому.

Последовательный монтаж – когда эпизоды монтируются в хронологическом порядке.

Метод параллельного монтажа заключается в том, что два события, связанные между собой по сюжету, показываются попеременно, одно за другим.

Исходный монтаж. Профессионалы утверждают, что в первом кадре любого рекламного сообщения должно быть заложено все интригующее притяжение, иначе говоря, первый кадр должен быть символичен. Как правило, любой новый кадр привлекает внимание зрителя. Первая его реакция – понять, что он видит. Для этого, в зависимости от крупности и сложности построения кадра, ему необходимо 2–4 с. Если в статичном кадре ничего не происходит в течение 4–6 с, то интерес зрителя к нему пропадает. Вернуть внимание зрителя можно либо сменой кадра, либо увеличением длительности кадра более 10 с, когда зритель начинает всматриваться в детали, искать нюансы, осознавать смысловую нагрузку кадра, либо началом действия в кадре.

Обычно если первый кадр композиционно более сложный, то следующий за ним должен быть проще и короче. Единого ритма в фильме можно добиться путем сокращения кадров до минимально возможной длины.

Монтаж по масштабности. В зависимости от сценария применяют сверхкрупный план или деталь-кадр, крупный план-кадр, первый средний план – человек по пояс, второй средний план – человек по колени, общий план – человек в полный рост, дальний план. Наиболее эффектно воспринимается стык между планами, находящимися на приведенной выше шкале, через один, т.е. общий план монтируется с первым средним и, наоборот, второй средний – с крупным, и т.д.

Монтаж по ориентации в пространстве. Предположим, что велась съемка выступления рекламодателя. Место оператора в зале находилось недалеко от сцены, чуть правее центра, а рекламодатель во время выступления в основном смотрел в центр зала. На записи преобладающее направление его взгляда будет справа налево. В какой-то момент оператору захотелось поснимать публику, он повернулся направо и включил запись.

Если все так и было, то при видеомонтаже этих кадров можно столкнуться с серьезной проблемой. Дело в том, что и выступающий, и публика на этих кадрах смотрят в одну сторону, справа налево. Проблема бы не возникла, если бы для съемки публики оператор повернулся не направо, а налево. Тогда глаза исполнителя и людей в зале имели бы встречное направление, и было бы ясно, кто на кого смотрит.

Простейший случай видеомонтажа по положению объектов в пространстве – видеомонтаж диалога двух персонажей (так называемая "восьмерка"). Если взгляды этих людей будут направлены навстречу друг другу, кадры смонтируются, если в одну сторону – нет. Правило восьмерки: двух разговаривающих можно снимать, только расположившись с одной стороны от них.

Следует отметить, что горизонтальная композиция кадра в видео оправдана тем, что человеческое зрение панорамно в ширину, а не в высоту, ибо кино или видео по своему существу являются "видом из глаз наблюдателя".

Монтаж по направлению движения. Любое действие в кадре – это движение, и зрачок человеческого глаза инстинктивно притягивается к любому движущемуся объекту, который становится центром зрительского внимания в кадре. В статичном кадре центром внимания является центр композиции.

При съемке кадров во время движения по улицам через правое и левое окно обязательно возникнут трудности при монтаже, связанные с тем, что кадры, снятые через левое и правое окно, "не захотят" стоять рядом. Зрителю будет казаться, что эти кадры были сняты из автомобилей, движущихся навстречу друг другу. Чтобы выйти из положения, нужно во время съемки сделать несколько кадров через переднее стекло автобуса. Такие кадры будут монтироваться с любым из предыдущих кадров. Кроме того, если в конце кадра, снимаемого, например, через левое окно, перевести камеру на какой-нибудь неподвижный объект (крупный план человека, сидящего у окна), то следующий кадр смонтируется "без вопросов". Если на одном кадре объект движется от нас чуть налево, то в следующем кадре он не должен двигаться направо. Операторы всегда начинают и заканчивают съемку панорам короткими статичными кусками. Это дает возможность монтировать панорамы через небольшую паузу 15 движения.

Монтаж по фазе движения. Если в конце общего плана человек начал поднимать, левую руку, то в начале среднего плана эта рука также должна подниматься. Иначе изображение не "склеится". Фазу приходится учитывать при видеомонтаже циклически повторяющихся положений объекта. В рекламе многие эпизоды снимают несколько раз, чтобы иметь несколько дублей одного кадра. Одно и то же действие, как правило, снимается несколько раз планами разной крупности. Монтажер, таким образом, имеет возможность на видеомонтажном столе подгонять каждое движение по фазе с точностью до кадра.

Монтаж по композиции (смещение центра внимания). Если на общем плане человек заметно смещен в одну сторону кадра, а на среднем – в другую, при просмотре в месте склейки зритель на время потеряет из вида объект из-за резкого смещения центра внимания. Чтобы этого не произошло, во время съемки и при видеомонтаже нужно помнить, что смещение центра внимания по горизонтали при переходе от кадра к кадру не должно превышать одной трети ширины экрана.

Монтаж по свету. Соседние кадры не должны резко отличаться по тону и характеру освещения. Это относится и к ситуации, когда, например, дальний план снят при солнечном освещении, а общий или средний – при пасмурном. Профессионалы в таких случаях снимают облака в тот момент, когда они закрывают солнце. Такой промежуточный видеомонтажный кадр позволяет перейти от "солнечного" кадра к "пасмурному".

Монтаж по цвету. Соседние кадры в месте стыка не должны резко отличаться по цвету. Цветовое решение соединяемых кадров в комфортном монтаже не должно сталкивать контрастные цвета – здесь действуют законы колористики: соседствующие цвета радуги "красный – оранжевый – желтый – зеленый – голубой – синий – фиолетовый" сочетаются, "зеленый" не только делит цвета на две гаммы – "красную" и "синюю" (которые несоединимы между собой), но и сочетается только со своими "соседями". Переход от одного основного цвета кадра к другому за одну склейку невозможен, необходимо поэтапное изменение соотношений цветов в несколько шагов. Если в новом кадре возникают новые цвета, то они должны занимать не более одной трети площади кадра.

"Перебивка". Перебивка – это кадр, который вклеивается между двумя другими кад-

рами, связанными между собой единством объектов и места действия. Содержание перебивки всегда резко отличается от предыдущего и следующего за ней кадров, но оно должно быть прямо или косвенно связанным с основным содержанием. Наиболее часто перебивки применяются при видеомонтаже длинных монологов. Если нужно сократить часть выступления человека, снятого длинным статичным средним планом, то без перебивки не обойтись. Для вставки перебивки выбирается место примерно за 1–2 с до того, когда говорящий делает небольшую паузу. К среднему плану клеится перебивка, во время которой звучит конец фразы. Следующий кадр – опять средний план героя, который начинает говорить с нужного редактору места. "Перебивкой" в этом случае может служить кадр, на котором снята картина, висящая на стене кабинета, часы или книга на письменном столе и т.п.

Звук в видео. Запись фонограммы сопряжена со многими трудностями, она требует изобретательности, тонкого слуха и развитого художественного вкуса. Только при соблюдении этих условий достигается та убедительность, к которой, как правило, стремятся режиссеры.

Качественная запись звука – это прежде всего ясность, четкость и естественная тональность звучания, которые достигаются путем грамотного размещения микрофонов. Современные микрофоны обладают одним приемным каналом и не пригодны к избирательной передаче звуков, а человек слышит ушами и способен воспринимать звуки избирательно. Поэтому микрофон нужно ставить в такое положение, чтобы он передавал на записывающую аппаратуру только необходимые звуки. Посторонние шумы, не связанные с содержанием кадра, источник которых находится за его пределами, способны запутать зрителей.

При съемке для контроля звука следует пользоваться наушниками. Если в наушниках слышен ветер, то надо поискать защищенное место или экранировать микрофон – хотя бы головным убором. Желательно микрофон камеры снабдить меховым чехлом для защиты от ветра. У некоторых камер есть ветрозащита, которая подавляет низкие частоты, что иногда дает хороший эффект.

Дополнительным может служить только направленный микрофон. Чем лучше его направленность, тем меньше записывается помех. Направленные микрофоны имеют трубчатую конструкцию, которая обеспечивает задержку низких частот. Для внешних микрофонов изготавливается иногда ветрофильтр, подавляющий более широкий диапазон низких частот. Направленные микрофоны могут монтироваться на камкордере, и ассистент не будет нужен. Для интервью и выступлений этот микрофон весьма полезен, так как обычно камера находится слишком далеко, чтобы хорошо записать текст.

Если микрофон будет стоять слишком близко к источнику звука, запись окажется лишеной глубины, будет восприниматься как зажата и грубая. Если же микрофон поставить на большом расстоянии от источника звука, запись получится глухой или с выделением отдельных случайных частот, что нежелательно. Если надо записать речь одного человека в студии или в комнате, микрофон устанавливают на уровне губ говорящего и, как правило, не ближе 60 см.

Рекламные фильмы сопровождаются музыкой и речью диктора-комментатора. Когда невозможно заранее предусмотреть содержание снимаемого материала, режиссерский сценарий не включает дикторского текста. Он пишется позднее на основании монтажных листов, в которых описаны содержание и продолжительность отснятых эпизодов. Опытный автор дикторского текста хорошо знает, какое количество слов возможно произнести за ту или иную единицу времени.

Фильм с минимумом авторского текста, как правило, обеспечен фонограммой, которая включает шумы, музыку, звуки природы. Свист ветра, стук колес мчащегося поезда, цокот копыт, звуки уличного движения могут быть отдельно записаны или взяты из специальной библиотеки шумов.

Речь диктора хорошо сочетается с музыкальным сопровождением. Музыка подбирается заблаговременно в соответствии с темой и жанром фильма, ее темп, ритм, настроение должны сочетаться с изображением и влиять на восприятие произведения. Неумелое музы-

кальное оформление может исказить смысл сюжета, побудить зрителя к неправильному истолкованию его содержания. Подбор и характер звучания музыкальных инструментов также оказывают воздействие на зрительское восприятие.

Подготовка монтажных листов. Прежде чем начать видео-монтаж, нужно провести некоторую подготовительную работу – написать видеомонтажные листы исходного материала. Если есть возможность вывести на экран телевизора показания счетчика ленты – очень хорошо, это существенно облегчит процесс. На этом этапе необходимо просмотреть весь исходный материал и выписать содержание и показания счетчика (начало и конец) тех кадров, которые собираетесь использовать в своем рекламном фильме. На этом этапе можно опускать бракованные и малоинтересные кадры. Если материал, из которого вы хотите монтировать фильм, располагается на двух или более кассетах, они должны быть пронумерованы, а их номера вписаны в видеомонтажные листы, чтобы не было путаницы, вызванной тем, что кадры из разных кассет будут иметь одинаковые показания счетчика. Если на исходной кассете есть "дыра", т.е. где-нибудь в середине показания счетчика прерываются и опять начинаются с нуля, это также необходимо отметить в видеомонтажном листе.

Монтажный план фильма. После того как были подготовлены видеомонтажные листы исходного материала, можно выбирать из видеомонтажных листов кадры в той последовательности, которая является наиболее подходящей.

Для того чтобы происходящее на экране было понятным и не затянутым по времени, в кино- и видеомонтаже используется прием – чередование крупности. Чтобы зрителю было понятно, где происходит действие фильма или отдельного эпизода, очень желательно, чтобы вначале в них присутствовали кадры, где объект снят общим планом. Это может быть один или несколько статичных кадров или кадров с панорамами или наездами. Злоупотреблять общими планами не стоит, поскольку, давая представление об объекте или действии в целом, они не позволяют разглядеть интересные подробности, которые чаще всего просматриваются на средних и крупных планах. Средняя длина общих планов – 5–8 с. Длина кадра, конечно, зависит и от характера объекта – если картинка лаконичная, можно обойтись тремя секундами, если же на ней много интересных деталей, которые хочется рассмотреть, то можно подержать общий план и подольше.

Рекламными образами могут быть люди, звери, растения. Объект рекламирования должен заинтересовать зрителя либо своим внешним видом, либо тем, что и как он делает, либо тем и другим одновременно. При этом необходимо избегать длинных видеомонтажных кадров, показывая объект в действии. В дополнение к этому можно вставить "перебивку", которая может содержать изображение, не относящееся напрямую к действию, но помогающее соединить между собой два нужных кадра.

Параллельный видеомонтаж – еще один способ видео-монтажного сокращения времени. Он используется при параллельном показе сюжета, растянутого во времени.

О спецэффектах-фэйдерах. Фэйдер – функция видеокамеры, позволяющая выполнять переход в затемнение и выход из затемнения. Лучшее место для применения этого эффекта – разделение эпизодов.

Визуальные эффекты – это специальные методы, используемые для создания и комбинирования несуществующих действий и сцен, получая в результате реалистичное изображение. Они применяются тогда, когда реальная реализация этих сцен слишком дорогостоящая или опасная, или попросту подобное невозможно снять вживую. Стоит обратить внимание, что «спецэффекты» и визуальные эффекты – это разные понятия. К спецэффектам относятся и всевозможные атмосферные эффекты на съёмочной площадке – дождь, снег, туман, лёгкое задымление, – и различные пиротехнические эффекты типа взрывов и попадания пуль, когда закладываются маленькие заряды и взрываются в определённое время, в то время как визуальные эффекты – это эффекты полностью выполненные на компьютере с помощью специальных программ. Они включают в себя набор различных технологий, методов и приемов. Одни из них используются во время съёмочного процесса, в то время как другие – в период последующей обработки.

Целью данной работы является ознакомление с разнообразными визуальными и спецэффектами, а также технологиями создания и применения их в мультимедийных продуктах, таких как видеоролики, рекламные клипы, фильмы. В данной работе были рассмотрены методы создания визуальных эффектов в компьютерной графике.

Специалист, занимающийся данными эффектами, называется VFX-artist (visual effects artist) или же художник по спецэффектам.

В середине прошлого века в основе создания спецэффектов были такие приемы, как ускоренная, замедленная и обратная киносъемка, рирпроекция и двойная экспозиция, стоп-камера и т.д. Именно благодаря этим эффектам, появились вполне реалистичные киноленты о чудовищах, таинственных превращениях и исчезновениях и т.д. Современная 3D-графика помогает в тех случаях, когда требуется построить воображаемую сцену в изображение реального мира. Такой способ называется композитинг – создание целостного изображения путём совмещения двух и более видео-слоёв материала, а также компьютерной графики и анимации. Такая ситуация типична для рекламных роликов, музыкальных клипов, кино- и видеоэффектов.

В процессе развития телевидения и кинематографа появились новые технологии, а некоторые из уже существующих устаревают и заменяются другими.

Причем каждая технология или прием имеют свое обозначение, которым оперируют специалисты в профессиональной практике. Набор обозначений всех технологий визуальных эффектов классифицируют по методу изготовления, а разделив все множество технологий на подмножества по моменту использования, их можно сгруппировать в некоторые классы. На сегодняшний день в профессиональной литературе визуальные эффекты делят на специальные и визуальные эффекты, а визуальные – на оптические и цифровые.

Есть более современные варианты классификаций, которые так же появились под влиянием попыток разделить визуальные эффекты на категории, объединяющие их по тем или иным признакам производства.

Например, классификация, которая приводится А. Кэрлоу, специалистом по производству визуальных эффектов: «В целом все спецэффекты можно разделить на четыре большие семейства, однако некоторые специфичные методы обладают перекрестными характеристиками.

Некоторые спецэффекты основаны на согласовании исходного «живого» действия: таков, например, метод камерного согласования. Другие предусматривают комбинирование элементов, полученных из нескольких источников (композитинг), добавление или удаление элементов, а также преобразование исходного изображения (морфинг)».

Сюда же относится классификация, приведенная Степановой М.А., которая делит визуальные эффекты на четыре категории: механические, оптические, грим и цифровые. Подобная классификация удобна для специалистов как рабочий инструмент, но она не может являться главным инструментом при анализе художественной природы визуальных эффектов и их роли в формировании экранного образа.

VFX-художник делает специальные эффекты для фильмов или рекламных роликов, его работа очень разнообразна: один день он может делать выстрелы и взрывы для экшн фильма, а на следующий симуляцию воды для синематика.

Например, в фильме «Годзилла» (реж. Р. Эммерих, 1998), Годзилла идет по улице и зритель видит только хвост монстра, который задевает стену здания, стена рушится, хвост скрывается за поворотом, люди в панике кидаются в стороны.

Конечно, этот кадр комбинированный, а Годзилла – трехмерный персонаж.

Падающие камни сделаны в трехмерном пакете с использованием модулей физических взаимодействий. И хвост, и камни введены в кадр с помощью компьютерной технологии, называемой композитинг. Камера при съемке двигалась.

А для того чтобы и хвост, и камни повторяли это движение, использовались или технология трехмерного трекинга или контроль движения камеры. На переднем плане пробегает человек, он перекрывает своим телом всю сцену, это означает, что человек вырезан маской (или снят на хромакее) и наложен поверх всей сцены.

Не трудно заметить, что при создании такой небольшой сцены перемещения монстра по городу были использованы следующие визуальные трехмерная графика, композитинг,

трекинг, вырезание масок, хромакей, моушн контроль и еще многие другие. Как правило, основной набор наиболее популярных цифровых технологий включает в себя следующие.

1 Композитинг представляет собой объединения нескольких слоев материала в один с помощью масок, полученных различными путями. Каждый слой может представлять собой: съемочный материал, фотографии, трехмерную или двумерную компьютерную графику. Композитинг является конечной операцией при создании конечного изображения.

2 Клинап. Описанный выше композитинг лежит в основе технологии клинапа «заплата», которая используется как для удаления объектов или изображения из кадра, так и для добавления в кадр новых объектов. И в том и в другом случае осуществляется одна и та же операция наложения на исходное изображение фрагмента другого, заранее заготовленного изображения. Так, чтобы убрать рекламу со стены здания, необходимо эту часть изображения перекрыть заготовкой фрагмента стены.

3 Маски. При композитинге совмещение всех слоев осуществляется с помощью масок, которые позволяют выделить в слое тот или иной объект и убрать все остальные части изображения. Получить маску можно различными путями. Наиболее простой, но и самый трудоемкий путь создания маски – вырезание вручную.

4 Моушнкепчер. Эта технология используется для переноса движений с тела реального актера на фигуру персонажа, созданного средствами трехмерной графики, в процессе его анимации. Различают моушнкепчер лица и тела.

5 Хромакей – технология очень похожая на люмокей, но отличается тем, что съемки ведутся не на черном фоне, а на цветном.

6 Люмокей – технология получения маски за счет съемки на черном фоне. Применяется для съемок различных сред и субстанций (пламя, вода, пыль, дым).

7 Кеинг – это компьютерная технология, которая используется для выделения объектов, снятых с помощью люмокея и хромакея, с процессе композитинга.

8 Морфинг позволяет осуществить плавное превращение одного изображения в другое путем его постепенной непрерывной деформации. В результате плавных трансформаций форма утрачивает определенность, становится текучей, тягучей, оплазмированной.

9 Ретайминг. С помощью ретайминга можно регулировать скорость демонстрации материала в кадре, искусственно замедляя или ускоряя его.

10 Трекинг позволяет восстановить на основе данных, содержащихся в съемочном материале, все параметры движения и изменения фокусного расстояния камеры. Используя трекинг в кадре можно добавить объекты при любом, самом сложном движении камеры.

Современная трехмерная графика позволяет создавать цифровое пространство и наполнять его объектами, моделируя законы реального мира. Основной набор инструментов: трехмерные объекты, виртуальные источники света и камеры, системы частиц и прочее. Все элементы трехмерной графики могут быть анимированны. Одним из простых эффектов является – метод анимированных спрайтов. Спрайт – это участок плоскости, который все время отображается перпендикулярно линии "взгляда" камеры.

Технология очень проста: создается анимированный спрайт, а в качестве текстуры накладывается последовательность кадров того или иного эффекта.

Все трехмерные спецэффекты можно разделить на несколько категорий:

- симуляции частиц (particle simulation);
- симуляция твердых тел (Rbs);
- симуляция мягких тел (Sbs);
- флюидные симуляции (Flip и Fluid simulation);
- симуляции толпы (crowd simulation).

Симуляции частиц – процесс настройки системы частиц.

Частицы, генерируемые компьютером, лежат в основе этого метода, суть которого состоит в моделировании движения частиц под действием различных сил. С помощью выполняемой разными способами визуализации этих частиц можно симулировать многочисленные

материалы и субстанции, такие как огонь, дым, туман, жидкости, грунт и даже волосы. Частицы – это элементарные спрайты, треугольники, маленькие полигоны или геометрические фигуры малых размеров (сферы, кубы и т.д.). Каждая частица в системе динамически меняет свои координаты в соответствии с заранее прописанным алгоритмом (например, по закону всемирного тяготения). Иногда движение частиц задается путем подкачки значений координат из внешнего файла, но часто большинство разработчиков просто прописывают один общий алгоритм движения частиц.

Симуляция твердых тел. Эта технология наиболее актуальна для разбиения объектов в результате столкновения или разрушения, но, в отличие от частиц, которые движутся только в трехмерном пространстве, и вычисляется по вектору, твердые тела занимают пространство и имеют физические свойства, такие как центр тяжести, момент инерции, и, самое главное, они могут иметь шесть степеней свободы (передвижение по всем трем осям, плюс вращение в трех направлениях).

Твердые тела представляют собой монолитные объекты, которые двигаются и вращаются в соответствии с назначенной им движущей силой.

Симуляция мягких тел – создание деформируемых объектов (поведение веревок, ткани, деформация машины, которая гнется под действием инерции при ударе и прочее). Есть много методов просчета данного вида динамики, но в последние годы все большую популярность завоевал FEM (finite element method) или метод конечного элемента, который ранее использовался только для инженерных расчетов, так как требовал огромных мощностей для вычислений.

Флюидные симуляции – симуляции таких эффектов как огонь, дым, взрывы, водные потоки.

Симуляция толпы – симуляция поведения огромного числа объектов с определенными похожими признаками по поведению, особенностям движения, например, толпы людей на стадионе, стаи птиц, косяки рыб.

Не трудно заметить, что за последние годы произошел качественный скачок – развитие компьютерных технологий вывело спецэффекты на новый уровень. Цифровые технологии открыли перед специалистами безграничный простор для фантазии. В данной работе приведены основные технологии создания трехмерных эффектов в компьютерной графике. Однако, спецэффекты – это всего лишь инструментарий, своего рода палитра для художника, они помогают раскрыть идею, сюжет фильма или видеоролика. Результаты данного исследования могут послужить основой для создания анимационных роликов и клипов, так как спецэффекты имеют прямое отношение к анимации, ибо принципы у них одинаковые – кадр за кадром работа с изображением.

Компьютерные технологии позволяют зрителю окунуться в неведомые миры, ощутить новую виртуальную реальность, и значение их не исчерпывается эффектом аттракциона и психологическим воздействием на зрительное восприятие, они меняют мышление, стимулируют фантазию и творчество.

РАЗДЕЛ 4. ПРОГРАММНЫЕ СРЕДСТВА ДЛЯ РЕАЛИЗАЦИЯ МУЛЬТИМЕДИЙНЫХ ПРОЕКТОВ

Тема 4.1. Средства разработки мультимедиа-приложений

Мультимедиа (multimedia) - это совокупность компьютерных технологий, одновременно использующих несколько информационных сред: графику, текст, видео, фотографию, анимацию, звуковые эффекты, звуковое сопровождение, человеческую речь.

Мультимедийные технологии - это совокупность современных цифровых средств аудио-, теле-, визуальных и виртуальных коммуникаций, которые позволяют вводить, сохранять, перерабатывать и воспроизводить текстовую, аудиовизуальную, графическую, трёхмерную и иную информацию.

Связывание элементов мультимедиа в единый проект выполняется с помощью про-

граммных средств. Результаты представления элементов мультимедиа на экране и средства управления мультимедиа, называются пользовательским интерфейсом, а аппаратные и программные средства, обеспечивающие воспроизведение мультимедиа, - платформой.

К разновидностям мультимедиа относятся:

Линейное мультимедиа - простейшая форма представления множества элементов мультимедиа, когда пользователь может выполнять только пассивный просмотр элементов мультимедиа, а последовательность просмотра элементов мультимедиа определяется сценарием.

Нелинейное (интерактивное) мультимедиа - форма представления множества элементов мультимедиа, в которой пользователю предоставлена возможность выбора и управления элементами в режиме диалога.

Гипермедиа - интерактивное мультимедиа, в котором пользователю предоставляется структура связанных элементов мультимедиа, которые он может последовательно выбирать.

Реальное видео - форма мультимедиа, моментально транслирующая поток данных с одного устройства на другое, предоставляющая пользователю просматривать видео и звуковое сопровождение в режиме реального времени.

В целом, под мультимедиа могут понимать и мультимедийную программу-оболочку, и продукт, сделанный на основе мультимедийной технологии, и компьютерное оснащение. Поскольку технологии мультимедиа являются комплексными, то и отдельные элементы этих технологий характеризуются многосредностью и ведением диалога с пользователем. Мультимедийные ресурсы, например, содержат различные виды информации, их существенной особенностью является активное взаимодействие ресурса и человека.

Технология мультимедиа является одной из новых технологических форм информационного общества. Она открывает принципиально новый уровень обработки информации и интерактивного взаимодействия человека с компьютером. Отличительной чертой современных мультимедийных технологий является их способность не только производить некий предназначенный для употребления продукт, но и оказывать косвенное влияние на пользующегося ими человека. Новые виды обработки и предоставления информации, новые способы доступа к информации позволяют разнообразить нашу культуру, содействуют глобальному обмену культурными ценностями, информацией и знаниями, способствуют более интенсивной коммуникации между людьми.

Исторической спецификой современной электронно-коммуникационной системы является то, что в отличие от прежних форм и стадий культурного развития человечества нынешняя характеризуется глобальными масштабами своего распространения и воздействия на все сферы общественной жизни.

Поскольку обмен информацией - необходимая составляющая жизни общества, то медиа технологии, как опосредующее звено человеческой деятельности, являются одним из способов коммуникации, условием человеческой активности. При этом интеграция в одной системе различных источников и форм информации в условиях открытого доступа фундаментально изменила характер коммуникации. Электронные цифровые медиаресурсы создают техническую возможность существования сверхнасыщенного информационного поля, которое практически повсеместно окружает современного человека.

Мультимедиа, помимо значительного ускорения коммуникативных процессов, позволяет на качественно новом уровне организовать процессы производства, хранения и распространения информации.

Активно внедряясь в деловую среду, мультимедиа влияет на ход экономического развития общества, рождая новое направление - электронный бизнес. Мультимедиа технологии широко используются в рекламной деятельности, при управлении маркетингом и организации продвижения товаров и услуг различными методами. Мультимедийные технологии становятся самостоятельным бизнесом и профессиональной областью деятельности, предметом бизнеса.

Невозможно переоценить значение мультимедиа в развитии индустрии развлечений,

создании компьютерных игр, киноиндустрии.

Мультимедиа следует рассматривать и как искусство, где особое место принадлежит наглядно-образным способам передачи информации. Как новая форма художественного творчества, мультимедиа выступает не столько продуктом технологической революции, сколько цифровым воплощением идей, которые не находили перспектив реализации в традиционных рамках изобразительного искусства и других видах культуры. При этом компьютер становится еще одним перспективным инструментом для всех искусств, альтернативной средой, способной по-новому реконструировать культуру и творить собственное искусство, он осознается как средство создания видов искусства. Сформировалось несколько направлений компьютерных искусств: цифровая музыка, интерактивный перформанс, компьютерная графика и анимация. Одним из основных преимуществ этих видов творчества считается открытость художественного пространства.

Одна из возможностей продуктивного использования мультимедиа - обучение. Мультимедийная технология позволяет увеличить степень усвояемости изучаемого материала, так как предоставляет возможность синергетического обучения. Под этим понимается обеспечение одновременно зрительного и слухового восприятия материала, активного участия в управлении его подачей, возвращения к тем разделам, которые требуют повторного анализа. Особенно велика роль мультимедиа технологий в развитии дистанционного образования. В будущем роль мультимедиа в области образования будет возрастать, так как знания, обеспечивающие высокий уровень профессиональной квалификации, всегда подвержены быстрым изменениям.

Существует большое множество программных средств для работы с мультимедиа файлами. Такие приложения можно разделить на несколько основных категорий:

- Средства создания и обработки изображения
- Средства создания и обработки 2D и 3D - графики
- Средства создания и обработки видео и анимации
- Средства создания и обработки звука
- Средства создания презентации

Проведя обзор и анализ существующих отечественных и зарубежных систем по технологии создания мультимедийных приложений, можно предложить следующую классификацию самых распространенных мультимедиа-приложений и их понятий. Мультимедийные приложения подразделяются на следующие виды:

- презентации;
- анимационные ролики;
- игры;
- видеоприложения;
- мультимедиа-галереи;
- аудиоприложения (проигрыватели звуковых файлов);
- приложения для web.

В табл. 1 представлены основные понятия мультимедийных приложений и их виды.

Таблица 1. Основные понятия мультимедийных приложений

Вид мультимедийного приложения	Понятие
Презентация	Презентация (от англ. presentation) – способ наглядного представления информации с использованием аудиовизуальных средств. Презентация представляет собой сочетание компьютерной анимации, графики, видео, музыки и звукового ряда, которые организованы в единую среду. Как правило, презентация имеет сюжет, сценарий и структуру, организованную для удобного восприятия информации
Анимационные ролики	Анимация – технология мультимедиа; воспроизведение последовательности картинок, создающее впечатление движущегося

	изображения. Эффект движущегося изображения возникает при частоте смены видеок кадров более 16 кадров в секунду
Игры	Игра – мультимедиа-приложение, направленное на удовлетворение потребностей в развлечении, удовольствии, на снятие напряжения, а также развитие определенных навыков и умений
Видеофильм и видео-проигрыватели	Видеофильмы – технология разработки и демонстрации движущихся изображений. Видеопроигрыватели – программы управления видеофильмами
Мультимедиа-галереи	Галереи – собрание изображений
Проигрыватели звуковых файлов (цифровой звук)	Проигрыватели звуковых файлов – программы, работающие с цифровым звуком. Цифровой звук – это способ представления электрического сигнала посредством дискретных численных значений его амплитуды
Приложения для web	Приложения для web – это отдельные веб-страницы, их компоненты (меню, навигация и т. п.), приложения для передачи данных, многоканальные приложения, чаты и т. д.

В свою очередь, мультимедийные приложения можно разделить на следующие подвиды. Основные понятия подвидов мультимедийных приложений представлены в табл. 2.

Таблица 2. Основные понятия подвидов мультимедиа-приложений

<p>Презентация:</p> <p>Линейная презентация – динамичный ролик со сложной графикой, видеовставками, звуковым сопровождением и отсутствием системы навигации.</p> <p>Интерактивная презентация – совокупность мультимедийных компонентов, структурированных по иерархическому принципу и управляемых через специальный пользовательский интерфейс.</p>
<p>Анимация:</p> <p>Покадровая анимация – кадровая смена изображений, создающая впечатление движения картинок.</p> <p>Программная анимация – анимация, при которой изображения меняются с помощью запрограммированной последовательности действий (то есть с помощью алгоритма и переменных). Рисование основных объектов происходит вручную, или импортирование их из коллекций и галерей, после чего применяются возможности какого-либо языка программирования.</p>
<p>Игры:</p> <p>Развлекательные игры – программы, позволяющие пользователю провести свой досуг.</p> <p>Обучающие игры – программы, позволяющие пользователю повысить уровень своих знаний в той или иной области, представленные в легкой игровой форме.</p>
<p>Видеопроигрыватели:</p> <p>Формирование покадрового фильма – подготовка и расположение изображений, последовательности фотографий, кадров, которые создают впечатление движения.</p> <p>Видеопроигрыватель для потокового видео – формирование проигрывателя, в который включается потоковое видео форматов avi, mpg и др., после чего появляется возможность управления этим потоком (например, использование таких команд, как запуск, пауза и перемотка на начало видеофрагмента).</p>
<p>Мультимедиа-галереи:</p> <p>Кадровая смена изображений – порядок смены изображений через определенный интервал времени.</p> <p>Панорама – широкая и многоплановая перспектива, позволяющая свободно обозревать большое открытое пространство.</p> <p>Интерактивная галерея – галерея, имеющая возможность управления пользователем (навигация по изображениям).</p>

Звуковые проигрыватели:

Проигрыватель одного звукового файла – добавление в мультимедиа-приложения звукового файла форматов wav, mp3 и др. и его воспроизведение.

Проигрыватель файлов разных звуковых файлов – аналогично проигрывателю одного звукового файла, но добавление такой возможности, как переключение между последовательностью исполнения.

Виртуальные музыкальные инструменты – имитация реальных музыкальных инструментов.

Приложения для web:

Баннеры – в Интернете, графическое изображение или текстовый блок рекламного характера, являющийся гиперссылкой на веб-страницу с расширенным описанием продукта или услуги. Баннеры размещают на веб-страницах для привлечения посетителей (потенциальных клиентов) или для формирования имиджа.

Приложения для передачи данных (например, гостевая книга).

Компьютерное представление графической информации реализуется с помощью растрового или векторного подхода. В первом случае изображение делится на пиксели, цвет каждого пикселя кодируется определенным числом битов. Векторные изображения сохраняются в виде геометрического описания объектов, составляющих рисунок.

Графические редакторы ориентированы на манипулирование существующими изображениями и обладают набором инструментов, позволяющих корректировать любой аспект изображения. Профессиональные графические редакторы поддерживают работу со слоями и экспорт объектов из программ векторной графики, обладают полным набором инструментов для коррекции цвета, ретуширования, регулировки контрастности и насыщенности цветов, маскирования, создания различных цветовых эффектов, имитирующих определенные художественные техники.

В программах векторной графики объекты и изображения, которые сохраняются в виде геометрического описания, существуют независимо друг от друга, что позволяет в любой момент изменять слой, расположение и любые другие атрибуты объекта, создавая произвольную композицию. В таких программах иллюстрации создаются с помощью фигур произвольной формы, их масштабирования, вращения, деформации, а также степени прозрачности и цветовой заливки. Современные программы векторной графики содержат также инструменты для работы с растровыми изображениями и текстами.

Трехмерная графика реализуется путем создания каркасов объектов, определения обтягивающих их материалов, компоновки всех объектов в единую сцену, установки освещения и точку визуализации – камеру. Для трехмерной анимации необходимо настроить перемещения объектов сцены и задать количество кадров. Движение объектов в трехмерном пространстве задается по траекториям, ключевым кадрам и с помощью формул, связывающих движение частей сложных конструкций. После задания нужного движения, освещения и материалов запускается процесс визуализации, в ходе которого просчитываются характеристики всех объектов сцены и генерируется последовательность изображений. Двухмерная анимация также использует традиционный покадровый принцип, только для создания последовательности используются двумерные изображения.

Для редактирования видео существует большое количество программных продуктов. Профессиональные видео-редакторы позволяют редактировать несколько видео- и звуковых каналов и осуществлять монтаж видеофрагментов в единую композицию. Они содержат наборы переходов между кадрами, синхронизируют звук и изображение, а также поддерживают редактирование и сохранение наиболее популярных форматов видеофайлов.

Программы для работы со звуком можно условно разделить на две большие группы: звуковые редакторы, ориентированные на цифровые технологии записи звука, и программы-секвенсоры.

Секвенсоры предназначены для создания музыки, с их помощью выполняется коди-

ровка музыкальных композиций, они используются для аранжировки, позволяя прописывать отдельные партии, назначать тембры инструментов, выстраивать уровни и балансы каналов, вводить музыкальные штрихи. Звуковые редакторы позволяют записывать звук в режиме реального времени на жесткий диск компьютера и преобразовывать его, используя возможности цифровой обработки звуковых частот и объединения различных каналов.

Средства создания презентаций, первоначально предназначенные для создания электронных слайдов, помогающих иллюстрировать сообщение докладчика, теперь все более ориентируются на применение мультимедиа. Существует большое количество таких программ, различающихся набором изобразительных и анимационных эффектов, способов управления презентацией и набором поддерживаемых мультимедиа файлов для импорта в качестве содержимого слайдов. По сути, презентация является информационным продуктом, объединяющим все мультимедиа форматы в одно целое.

Перспективы мультимедиа разнообразны, области применения будут расширяться, в том числе, благодаря появлению новых информационных технологий и способов обработки информации. Грамотное сочетание мультимедиа с другими технологиями будет способствовать более динамичному их развитию и еще большей интеграции во все сферы общества.

Для того, чтобы реализовывать эффектные и полезные мультимедиа приложения с использованием технологии Intel Perceptual Computing, в первую очередь, необходимо четко определить, какие именно мультимедиа форматы и в какой степени будут задействованы в конкретном приложении и какие технологии нужны для работы с этими мультимедиа форматами. Успех таких приложений будет в большей степени зависеть от того, насколько разработчики опыты в использовании тех или иных подключаемых библиотек функций, с помощью которых происходит обработка мультимедиа потока. Такое приложение будет полезно пользователю только в том случае, если оно корректно обрабатывает мультимедиа содержимое и является правильно реализованным с технической точки зрения. Поэтому, прежде чем начинать разработку мультимедиа приложения с использованием Intel Perceptual Computing, необходимо детально изучить принципы и особенности технологии обработки мультимедиа потоков, библиотеки функций и примеры программ. Только при условии, что управление с помощью технологии Intel Perceptual Computing будет дополнять полезный и корректно работающий мультимедиа-функционал, разработчику удастся воплотить все преимущества от синергии мультимедиа и Perceptual Computing.

Тема 4.2. Создание мультимедийной презентации

В деловой жизни часто приходится сталкиваться с ситуацией, когда необходимо заинтересовать людей своими достижениями, привлечь внимание к деятельности фирмы, рассказать в доступной форме о товаре, сделать доклад на конференции и пр.

В подобных случаях неоценимую помощь может оказать программа подготовки презентации. Главное – это придумать сценарий презентации, построенный на использовании картинок, анимации, броского запоминающегося текста.

В широком смысле слова термин презентация (от англ. *present* – представлять) – это выступление, доклад, защита перспективного или законченного проекта, представление рабочего плана, технического предложения, готового товара или услуги, результатов внедрения, контроля, испытаний и многое другое.

Презентация – это набор слайдов, связанных между собой возможностью перехода от одного слайда к другому и хранящихся в общем файле.

Подобно тому, как текстовый документ состоит из страниц, файл презентации состоит из последовательности кадров, или слайдов.

Слайд – логически автономная информационная структура, содержащая все объекты (элементы), которые представляются на экране (на цветной пленке, листе бумаги) в виде единой композиции. **Слайд** – это одна страница визуального материала вне зависимости от того, куда она будет направлена – на экран дисплея, принтер или фотопленку.

Слайды, создаваемые для электронной презентации, могут содержать текст, диаграм-

мы, рисованные объекты и фигуры, а также картинки, слайд-фильмы, звуки и графику, созданные в других приложениях. В электронную презентацию можно вносить изменения в последний момент; темп презентации регулируется *установкой* интервалов показа слайдов, а также использованием специальных переходов при смене слайдов и анимации. Электронную презентацию можно запустить в автономном режиме. Презентационную конференцию можно провести в сети на нескольких компьютерах.

Презентацию можно подготовить с расчетом ее эффектного показа как на экране в цвете, так и на бумаге или на прозрачной пленке, т.е. в виде материалов, распечатанных на лазерном принтере (в оттенках серой шкалы или в черно-белом виде). Перед печатью возможен предварительный просмотр презентации, а также внесение изменений. Для облегчения проведения презентации присутствующим можно представить раздаточный материал - печатный вариант презентации, содержащий по два, по три или по шесть слайдов на странице. Кроме того, для зрителей можно распечатать заметки докладчика.

Презентацию можно оформить специально для сети Web, а затем сохранить ее в одном из Web-совместимых форматов, например в формате HTML.

Прежде чем создавать презентацию необходимо тщательно продумать тему и сюжетный ход презентации, иначе говоря, разработать **сценарий** презентации.

Создание любой презентации предусматривает решение двух основных задач: **разработка содержания** презентации и **проектирование ее дизайна** (оформление презентации).

В презентацию могут входить элементы:

- шаблоны слайдов;
- плавные переходы между слайдами;
- звуковые эффекты при наступлении некоторых событий;
- анимация текста и изображений;
- набор изображений;
- авторские заметки.

В составе слайда могут присутствовать следующие объекты:

- заголовок (текстовый) и подзаголовок;
- графические изображения (рисунки);
- таблицы;
- диаграммы (графики);
- организационные диаграммы (оргдиаграммы);
- текст;
- звуки;
- маркированные списки;
- фон;
- номер слайда;
- дата;
- носители гиперссылок (объекты в составе слайда, которые инициируют переход к другому слайду, к другой презентации или к web-странице);
- различные внешние объекты.

Классификация презентаций

1. *HTML-презентации:*

- представляет собой самостоятельный сайт;
- может публиковаться как в Интернет, так и на компакт диске;
- при наличии адаптированной навигации может использоваться докладчиком при выступлении;

2. *MS PowerPoint-презентация:*

- разрабатывается с использованием уникальных графических элементов, создаваемых в соответствии с фирменным стилем компании;
- как правило, используется при публичных выступлениях или на выставках, с це-

лью усилить эффектность выступления и сделать предлагаемый вниманию слушателей материал более наглядным.

3. *Flash-презентация:*

- разрабатывается с использованием уникальных графических элементов, создаваемых в соответствии с фирменным стилем компании;
- может публиковаться как в сети Интернет, так и на компакт диске;
- предназначена как для использования в ходе выступления, так и для распространения на компакт дисках.

4. *Презентация смешанного типа:*

- использует одновременно несколько указанных выше технологий, объединенных, единым интерфейсом (программной оболочкой для просмотра);
- как правило, используются для публикации и распространения на компакт-дисках.

Презентации разрабатываются для достижения различных целей и выполняют разнообразнейшие функции от представления организаций и выпускаемой фирмой продукции до интерактивных видео экскурсий и различных подарочных изданий.

Презентации могут быть самых различных видов, например:

- представляющие государственные учреждения и коммерческие организаций;
- CD-визитки;
- каталоги в электронном виде;
- мультимедиа-руководства для обучения;
- личные архивы - электронных копий документов с возможностью интерактивного просмотра с текстовыми и голосовыми комментариями;
- представляющие дипломные работы, различного рода диссертации;
- поздравления;
- подарочные издания.

По способу представления информации мультимедиа презентации могут быть линейными и нелинейными. **Линейные** презентации создаются по написанному сценарию и демонстрируются пользователю в виде видео или мультфильмов без возможности управления показом. **Нелинейные** презентации предоставляют пользователю возможности управления показом, используя системы гипертекста.

Создание презентации в большинстве случаев включает четыре основных этапа.

1 этап. Сначала необходимо спланировать общий вид презентации, при этом программа должна напомнить, какие этапы целесообразно включить в создание презентации выбранного типа.

2 этап. Затем следует отредактировать тексты презентации и поместить рисунки. На этом этапе программа должна предоставить удобные средства редактирования текстов и рисунков. Кроме того, современные презентации должны позволять задавать специальные эффекты, определяющие тип демонстрации. Появление каждого нового слайда может сопровождаться тем или иным анимационным эффектом. Например, слайд может растворяться на экране, «вылетать» с него или даже распадаться на части.

3 этап. Естественно, программа должна обеспечивать вывод презентации на печать.

4 этап. Это этап — «упаковка». Она необходима в случае переноса презентации на мобильный компьютер для транспортировки к месту проведения презентации.

Существует много программ по созданию презентаций.

DarkBASIC - это уникальная среда программирования, позволяющая создавать собственные презентации, слайд-шоу и даже бизнес-приложения с использованием простейшего языка программирования BASIC. Интуитивный интерфейс, мощная интегрированная система справки и многочисленные примеры кода значительно облегчают работу в этой среде.

Программа **Keynote** - это незаменимый инструмент каждого делового человека, которому часто приходится публично представлять свои идеи. Keynote позволяет создавать профессионально оформленные презентации.

Программная оболочка **Formula Graphics** является одной из наиболее удачных сред

для создания презентационной продукции. Она имеет легкий в использовании графический интерфейс и не несет ограничений на изображения, звуки и анимации, которые могут быть объединены с ее помощью. Formula Graphics имеет мощный объектно-ориентированный язык с более чем 500 операторами, функциями и командами, но можно создавать приложения и без программирования. Подготовка данных производится в других программах, а Formula Graphics объединяет их в единое целое.

Презентационная программа есть в офисном пакете - WordPerfect Office компании Corel. Эта программа называется **Corel Presentation**.

Multimedia Builder MP3 – визуализированный программный продукт для создания мультимедийных презентаций.

MICROSOFT POWERPOINT представляет собой программу, позволяющую создавать файлы презентаций, включающие демонстрационные слайды, структуру, заметки докладчика и раздаточные материалы (например, презентации учебных курсов, бизнес-планов, деловых проектов или отчетов).

Программа PowerPoint создает мультимедийные презентации в виде связанных страниц - **слайдов**. В слайды можно вставлять текст, картинки, видео фрагменты, диаграммы и т.д. Кроме того, слайды могут сопровождаться музыкой или записанными звуковыми комментариями.

Мультимедиа-презентация - это программа, которая может содержать текстовые материалы, фотографии, рисунки, звуковое сопровождение, видеофрагменты и анимацию, трехмерную графику. Благодаря комбинации этих средств она является практически самой эффективной формой представления информации.

Информация в мультимедиа-презентации может быть организована различными способами:

1. Интерактивный проект с меню и расположением информационных страниц в иерархическом порядке.
2. Автоматически проигрывающийся ролик с повествованием (как правило, для удобства управления размещается навигационная панель).
3. Комплекс из интерактивных меню, информационных страниц и автоматически проигрываемых роликов.

Хотя приложение PowerPoint обладает собственными средствами для создания объектов различного типа (текст, таблицы, графики и т.д.), благодаря тесной интеграции с другими компонентами пакета Microsoft Office пользователь имеет возможность применять уже наработанные материалы. Например, текст может быть подготовлен в текстовом редакторе Word, формулы – в приложении Microsoft Equation, таблицы – в табличном процессоре Microsoft Excel, диаграммы – в приложении Microsoft Graph, художественные заголовки – в приложении Microsoft WordArt и так далее. Конечно, можно использовать некоторые специализированные приложения: Adobe Photoshop или CorelDraw, однако именно приложения, входящие в состав Microsoft Office, являются наиболее интегрированными и обмен данными происходит без потерь и искажений.

Для использования некоторых возможностей требуется дополнительное оборудование и программное обеспечение:

- при работе с мультимедиа и звуком для улучшенного отображения графики требуется видеоплата, поддерживающая ускорение графики, или процессор, поддерживающий набор команд MMX;
- микрофон ближнего действия и устройство воспроизведения звука;
- Microsoft Exchange, Internet SMTP/POP3, IMAP4 или другая MAPI-совместимая система для работы с электронной почтой;
- для организации совместной работы средствами Microsoft Outlook необходим сервер Microsoft Exchange Server;
- средства совместной работы, совместимые с пакетом программ Office 2000 или более поздней версии;

- для некоторых видов работы может потребоваться доступ к Интернету с внесением дополнительной платы поставщику Интернет-услуг, а также оплата услуг местной связи;
- модем со скоростью 14 400 бод или выше.

Профессиональные мультимедийные презентации пишутся с использованием языков программирования высокого уровня. К ним в первую очередь относится язык **Lingo**, интегрированный в мощную программу для создания презентаций **Macromedia Director**. В такой презентации одновременно может воспроизводиться несколько разнородных мультимедийных элементов - видео, аудио, flash, слайд-шоу - делая презентацию действительно современной и эффектной. Для просмотра презентации пользователю не придется совершать никаких дополнительных действий - установленная на компьютере Windows содержит все необходимые драйвера и плееры.

Процесс создания презентации в Microsoft PowerPoint состоит из таких действий, как выбор общего оформления, добавление новых слайдов и их содержимого, выбор разметки слайдов, изменение при необходимости оформления слайдов, изменение цветовой схемы, применение различных шаблонов оформления и создание таких эффектов, как эффекты анимации при демонстрации слайдов. Приведенные ниже сведения касаются средств, доступных на начальном этапе этого процесса.

PowerPlugs: Presentations to Go – это набор заготовок презентаций PowerPoint, позволяющий в считанные минуты создать первоклассное визуальное сопровождение для типовых встреч, совещаний и семинаров. Каждая заготовка содержит 12 профессионально оформленных слайдов, выполненных с учетом особенностей целевой аудитории и предмета обсуждения. Заготовки Presentations to Go снабжены профессиональными рекомендациями по организации и подаче информационного материала.

Способы создания презентации

- создание новой презентации с помощью пустых слайдов;
- создание новой презентации на основе существующей – презентация создается на основе уже имеющейся презентации с заданным оформлением; создается копия имеющейся презентации, позволяющая создать новую презентацию, внося изменения в оформление и содержимое исходной презентации;
- создание презентации с использованием шаблона оформления - презентация создается на основе имеющегося шаблона Microsoft PowerPoint, содержащего основные элементы оформления, шрифты и цветовую схему, кроме стандартных шаблонов Microsoft PowerPoint можно использовать самостоятельно созданные шаблоны; PowerPoint предлагает два типа шаблонов: шаблоны оформления и шаблоны содержания. Шаблоны содержания предлагают уже оформленные слайды с текстом на различные темы, которые можно изменять по своему усмотрению. Кроме того, на основе уже имеющейся презентации есть возможность создавать пользовательские шаблоны. Добавленный в **Мастер автосодержания** новый шаблон будет доступен при следующем использовании мастера. Шаблоны оформления предлагают разные стили форматирования и разное цветовое оформление отдельных слайдов.
- из мастера автосодержания – презентация создается на основе предложенного содержания (пошаговый метод создания новой презентации). Мастер автосодержания позволяет последовательно выбрать сначала **Вид презентации**, потом **Стиль презентации** и так далее. Предлагается около двадцати различных видов презентаций, в том числе «Общий доклад», «Бизнес-план», «Мозговой штурм», «Диплом» и др.

При создании презентации основное внимание необходимо уделять их содержанию. Анимация, переходы и другие инструментальные средства используются для подчеркивания определенных аспектов сообщаемых сведений, чтобы не отвлекать внимание аудитории на спецэффекты.

Подобных правил следует придерживаться и в отношении звука. Музыка, которая звучит во время перехода с одного слайда на другой или во время анимации, сконцентрирует внимание зрителей на показываемых слайдах. Однако злоупотребление звуковыми эффектами может отвлечь внимание слушателей от важных моментов.

Реакция аудитории зависит также от темпа проведения презентации. Так, слишком быстрая смена слайдов утомляет, а слишком медленная может подействовать расслабляюще. Средства программы *PowerPoint* позволяют отрететировать темп показа перед проведением презентации.

Можно предоставить общий доступ к презентации в Интернете, опубликовав ее на веб-сервере или другом компьютере, пользователям, которым требуется ознакомиться с презентацией с помощью веб-обозревателя. При опубликовании презентации ее копия помещается в выбранное местоположение. Опубликовать можно презентацию, сохраненную в формате.ppt, веб-страницы или веб-страницы в одном файле.

При публикации презентации в Интернете или сохранении ее в формате веб-страницы в нее автоматически включаются следующие элементы:

- Область переходов, отображающая структуру презентации.
- Область слайдов.
- Элемент управления, позволяющий отобразить или скрыть структуру презентации.
- Элемент управления для отображения и скрытия области заметок.
- Элемент управления, позволяющий включить режим просмотра во весь экран, аналогичный режиму демонстрации слайдов в Microsoft PowerPoint, при котором скрываются элементы управления обозревателем. При наличии в презентации произвольных показов их можно просмотреть только в полноэкранном режиме.

Опубликование презентации позволяет решить следующие задачи:

- предоставление общего доступ в Интернете к копии презентации, которую можно изменять и обновлять. Оригинал может храниться в формате презентации PowerPoint [.ppt];
- публикация копий исходной презентации на разных веб-узлах;
- предоставление доступа к определенной части презентации, например к произвольному показу, одному слайду или нужному диапазону слайдов;
- настройка презентации с учетом особенностей конкретной программы (или версии) веб-обозревателя, например Microsoft Internet Explorer или Netscape Navigator ;
- выбор элементов, отображаемых в веб-формате презентации, таких как заметки докладчика, анимации и кнопки перемещения.

Тема 4.3. Динамические процессы в мультимедиа-среде

Анимацией называется искусственное представление движения в кино, на телевидении или в компьютерной графике путем отображения последовательности рисунков или кадров с частотой, при которой обеспечивается целостное зрительное восприятие образов.

Анимация, в отличие от видео, использующего непрерывное движение, использует множество независимых рисунков.

Синоним «анимации» – «мультипликация» – очень широко распространен в нашей стране. Анимация и мультипликация – это лишь разные определения одного и того же вида искусства.

Более привычный для нас термин произошел от латинского слова «мульти» – много и соответствует традиционной технологии размножения рисунка, ведь для того, чтобы герой «ожил», нужно многократно повторить его движение: от 10 до 30 рисованных кадров в секунду.

Принятое в мире профессиональное определение «анимация» (в переводе с латинского «анима» – душа, «анимация» – оживление, одушевление) как нельзя более точно отражает все современные технические и художественные возможности анимационного кино, ведь мастера анимации не просто оживляют своих героев, а вкладывают в их создание частичку своей души.

В настоящее время существует различные технологии создания анимации:

Классическая (традиционная) анимация представляет собой поочередную смену ри-

сунков, каждый из которых нарисован отдельно. Это очень трудоемкий процесс, так как аниматорам приходится отдельно создавать каждый кадр.

Стоп-кадровая (кукольная) анимация. Размещенные в пространстве объекты фиксируются кадром, после чего их положение изменяется и вновь фиксируется.

Спрайтовая анимация реализуется при помощи языка программирования.

Морфинг – преобразование одного объекта в другой за счет генерации заданного количества промежуточных кадров.

Цветовая анимация – при ней изменяется лишь цвет, а не положение объекта.

3D-анимация создается при помощи специальных программ (например, 3D MAX). Картинки получаются путем визуализации сцены, а каждая сцена представляет собой набор объектов, источников света, текстур.

Захват движения (Motion Capture) – первое направление анимации, которое дает возможность передавать естественные, реалистичные движения в реальном времени. Датчики прикрепляются на живого актера в тех местах, которые будут приведены в соответствие с контрольными точками компьютерной модели для ввода и оцифровки движения. Координаты актера и его ориентация в пространстве передаются графической станции, и анимационные модели оживают.

Принципы анимации

При создании анимационных фильмов используются некоторые общие принципы. Большинство из них сформулировано для анимации Диснея и первоначально относилось к мультфильмам, выполненным в технике традиционной анимации, но практически все они применимы и при других технологиях. Вот основные из них:

«Сжатие и растяжение» (squash & stretch). Этот принцип произвел революцию в мире анимации. Суть принципа состоит в том, что живое тело всегда сжимается и растягивается во время движения. Перед прыжком персонаж сжимается как пружина, а в прыжке наоборот растянут. Главным правилом при этом является постоянный объем - если персонаж растянули (stretch - деформация по оси Y), то он обязательно должен быть сжат для сохранения объема своего тела (squash - деформация по оси X).

«Подготовительное действие» (Anticipation). В реальной жизни для произведения какого-либо действия, человеку часто приходится делать подготовительные движения. Например, перед прыжком человеку необходимо присесть, для того чтобы бросить что-либо руку необходимо завести назад. Такие действия называются отказными движениями, т.к. перед тем как сделать что-то персонаж как бы отказывается от действия. Такое движение подготавливает зрителя к последующему действию персонажа и придает инерцию движениям.

Сценичность (staging). Для правильного восприятия персонажа зрителями все его движения, позы и выражения лица должны быть предельно просты и выразительны. Этот принцип основан на главном правиле театра. Камера должна быть расположена так, чтобы зритель видел все движения персонажа.

«Ключевые кадры» (Pose to Pose). До открытия этого принципа движения рисовались, и поэтому результат было трудно предсказать, т.к. сам художник еще не знал, что он нарисует. Этот принцип предусматривает предварительную компоновку движений - художник рисует основные моменты и располагает персонажа на сцене, а уж потом ассистенты прорисовывают все кадры движения. Этот подход резко увеличил производительность, т.к. заранее планировались все движения, и результат был именно таким как задумывалось. Но чтобы создать какое-то конкретное движение, была необходима тщательная проработка каждого «кусочка». Разрабатывая выразительные позы художник вкладывает все свое мастерство, поэтому именно эти моменты должны быть дольше видны зрителю. Для этого ассистенты дорисовывают движения так, что больше всего кадров оказывается рядом с ключевыми позами. При этом персонаж как бы проскальзывает движение от одной компоновки к другой, медленно выходя из позы и замедляясь у другой.

«Сквозное движение и захлест» (follow through / Overlapping actions). Суть принципа

состоит в том, что движение никогда не должно прекращаться. Существуют такие элементы как уши, хвосты, одежда, которые постоянно должны находиться в движении. «Сквозное движение» обеспечивает непрерывность движения и плавность перехода фаз, например, из бега в шаг и наоборот. Движение отдельных элементов тела, в то время как тело уже не двигается, называется захлестом. Захлест выражается в сценах смены фаз движения. Если персонаж резко тормозит после бега, мягкие части тела не могут остановиться вместе с жесткими и происходит небольшой захлест (волосы, уши, хвосты и т. д). При ходьбе движение начинается с бедер, а уж потом распространяется до лодыжек. Таким образом, все движения персонажа связаны в отдельную цепочку, и появляется возможность жестко описать правила, по которым он двигается. Движение, при котором один элемент следует за другим, называется сквозным движением.

«Движения по дугам» (arcs). Живые организмы всегда передвигаются по дугообразным траекториям. До этого применялся метод прямолинейного движения, в связи с чем, движения выглядели механическими - как у роботов. Характер траектории зависит, как правило, от скорости движения. Если персонаж движется резко, траектория распрямляется, если же медленно, то траектория еще больше загибается.

Второстепенные действия (Secondary actions). Часто для придания персонажу большей выразительности используют вторичные движения. Они служат для того, чтобы акцентировать внимание на чем-нибудь. Например, горящий персонаж может часто сморкаться в платок, а удивленный подергивать плечами. Вторичные действия получили широкое распространение в мировой анимации. Благодаря их использованию персонажи становятся более живыми и эмоциональными.

Расчет времени (Timing). Этот принцип позволяет придать персонажу вес и настроение. Как зритель оценивает вес персонажей? Вес персонажа складывается из таких факторов как скорость перемещения и инертность. Для того чтобы персонаж двигался в соответствии со своим весом, художник рассчитывает время движения и захлеста для каждого персонажа. При расчете времени учитываются вес, инертность, объем и эмоциональное состояние героя. Настроение также передается скоростью движений персонажа. Так подавленный персонаж движется очень вяло, а воодушевленный достаточно энергично.

Преувеличение (Exaggerrate and Caricature). Уолт Дисней всегда требовал от своих работников большего реализма, на самом деле стремясь больше к "карикатурному реализму". Если персонаж должен был быть печальным, он требовал, чтобы его делали мрачным, счастливого же нужно было делать ослепительно сияющим. С помощью преувеличения увеличивается эмоциональное воздействие на зрителей, однако, персонаж приобретает карикатурный характер.

Профессиональный рисунок. Рисунок основа всего. На студии Диснея довольно часто встречаются таблички вроде: "Чувствуется ли в твоём рисунке вес, глубина и равновесие?". Принцип профессионального рисунка также воспрещает рисовать "близнецов". "Близнецами" называют любые элементы рисунка, которые повторяются дважды или являются симметричными "Близнецы" очень часто появляются помимо воли художника, сам не замечая того, он рисует две руки в одном и том же положении.

Привлекательность (Appeal). Привлекательность персонажа - путь к успеху всего фильма. Как же определить, привлекателен ли персонаж? Привлекательным может быть любой предмет, если смотришь на него с удовольствием, обнаруживая в нем простоту, обаяние, хороший дизайн, очарование и магнетизм. От привлекательного персонажа невозможно оторвать взгляд. Даже самый противный герой фильма должен быть привлекательным, чтобы удержать зрителей у экрана.

К анимированным относятся элементы типа Multistate Button и Multistate Level.

Multistate Button – анимированная кнопка, переход которой из одного состояния в другое сопровождается воспроизведением ряда кадров анимации, привязанных к ней. Анимированная кнопка отличается от обычной только визуально. Так же, на основе этого эле-

мента можно создать нефункциональную графику, проходящую с первого по последний кадр один или несколько раз, циклическую анимацию.

Multistate Level – анимированный слайдер, каждому положению шкалы которого присвоено собственное изображение (так же возможен вариант, что диапазон слайдера разделен на части, число которых соответствует числу кадров анимации, в этом случае кадр соответствует определенному диапазону значений). Так, например, можно реализовать изменение цвета шкалы с увеличением уровня (значения), движение стрелки и подобные эффекты.

Для получения анимированных элементов можно:

1. Использовать готовые элементы из библиотеки iRidium GUI Editor
2. Создавать собственные элементы с помощью инструмента массового импорта изображений в галерею. Таким способом можно создавать как обычные кнопки слайдеры, так и элементы с большим числом состояний.

В программах, ориентированных на создание анимированных объектов (таких как Macromedia flash), можно создавать анимацию, изменяя содержимое последовательности кадров. Анимлируемый объект можно заставить двигаться по сцене, увеличивать или уменьшать свой размер, поворачиваться, изменять цвет, исчезать или появляться, а также изменять форму. Все изменения могут происходить как независимо друг от друга, так и совместно. Существует два способа создания анимационных последовательностей во Flash: автоматическая (tweened) анимация и покадровая анимация:

Создавая автоматическую анимацию, достаточно задать начальный и конечный кадр, а Flash сам создаст промежуточные кадры. Flash самостоятельно будет изменять размеры объекта, угол поворота, цвет, и другие его параметры, распределяя эти изменения равномерно между начальным и конечным кадрами, создавая видимость движения.

Создавая покадровую анимацию, вы последовательно создаёте каждый кадр анимации

Можно использовать маскирующие слои для создания прозрачных масок, через которые видно содержимое одного или нескольких низлежащих слоёв. Также существует возможность создавать динамические маскирующие слои.

Инструменты большинства анимационных редакторов:

2. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ (УКАЗАНИЯ) К ЛАБОРАТОРНЫМ ЗАДАНИЯМ

Дисциплина «Мультимедиа-технологии» служит основой для знаний, умений и навыков взаимодействия с мультимедийной информацией. При подготовке к практическим работам предполагается изучение лекционного материала, а также теоретического материала из основной и дополнительной литературы.

Практическая работа считается выполненной, если предоставлен отчет о результатах выполнения задания; проведена защита проделанной работы.

Защита проводится в два этапа:

- демонстрируются результаты выполнения практического задания.
- ответить на ряд вопросов из перечня контрольных вопросов, который приводится в задании на практическую работу.

Каждая практическая работа оценивается определенным количеством баллов в соответствии с критериями оценивания.

Лабораторная работа № 1 Управление голосом. Анализ голоса. Сфера использования голосового управления. Прикладные задачи использования голосовых команд и интерактивного взаимодействия с устройствами.

Время выполнения - (4 ч)

Задание

- 1 Провести исследование мужского и женского голосов
2. Загрузить модель LPC-речевого кода.
3. Запустить процесс моделирования и зафиксировать осциллограммы и спектрограм-

мы входного, выходного сигналов, а также сигнала остатка предсказания. Прослушать входной и синтезированный сигналы. Сформулировать выводы относительно сравнения сигналов по форме и уровню, а также по результатам прослушивания.

4. Сохранить файл модели в своей папке. Прочитать в рабочее пространство массив заданного звукового файла. Провести моделирование процесса кодирования звукового файла с заданными параметрами: размером сегмента сигнала, величиной перекрытия сегментов, формой окна и количеством коэффициентов отражения (порядком предсказания). Оценить результаты моделирования согласно п. 4. При выполнении данного пункта задания следует использовать справочные данные по блокам моделей, приведенные в приложении.

5. Для заданного номера сегмента звукового файла построить график зависимости мощности остатка предсказания от порядка предсказания. Следует воспользоваться дополнительным выходом «Мощность остатка предсказания» блока **Levinson – Durbin**. Данный выход организуется путем дополнительной настройки блока. При выполнении данного пункта задания следует обозначить порядок предсказания через переменную, использовать операторы цикла, а также осуществлять запуск модели из командного окна.

Содержание отчета

Структурная схема модели.

Распечатка результатов моделирования.

Файлы *.mdl.

Контрольные вопросы

1. Как изменятся сигнал остатка предсказания и синтезированный сигнал при уменьшении порядка предсказания?

2. Имеются два кодека. В одном на вход синтезирующего фильтра подается сигнал остатка предсказания, а во втором – либо белый шум, либо последовательность импульсов. В каком кодеке уровень искажений сигнала больше?

3. Что следует изменить в модели кодека, чтобы увеличить громкость синтезированного сигнала?

4. Ответить на вопросы преподавателя по использованию блоков моделей.

Лабораторная работа № 2 Частотное редактирование звукового сигнала. Уровень громкости. Возможность стилизации звука при использовании звуковых эффектов.

Время выполнения - (4 ч)

Задание

1. Используя модель **LPC** кодека, в котором возбуждение синтезирующего фильтра осуществляется сигналом остатка предсказания, а также модель формирования импульсного сигнала возбуждения, создать модель кодека с импульсным сигналом возбуждения синтезирующего фильтра.

2. Прочитать в рабочее пространство заданный звуковой файл.

3. Используя модель **LPC** кодека, в котором возбуждение синтезирующего фильтра осуществляется сигналом остатка предсказания, определить с помощью функций **MATLAB** частоту основного тона в сигнале остатка предсказания для вокализованного сегмента речевого сигнала.

4. Установить в блоке импульсного генератора созданной модели период следования импульсов, соответствующий определенной ранее частоте основного тона. Рассчитать значение постоянной составляющей импульсной последовательности и установить полученное значение в блоке **Constant**.

5. Запустить созданную модель для заданного звукового файла.

7. Меняя период следования импульсов (одновременно меняя значение постоянной составляющей и параметра **Puls Width**), проследить за изменениями формы синтезированного сигнала и прослушать соответствующий звуковой файл.

8. Сформулировать выводы о различиях исходного и синтезированного речевых сигналов.

9. Создать модель кодека с комбинированным (импульсы – шум) сигналом возбуждения синтезирующего фильтра.

10. Повторить пункты 5-, 6-, 8-й для созданной модели. Подобрать оптимальное значение порога r классификации, обеспечивающее минимум искажений синтезированного сигнала.

Содержание отчета

1. Файлы *.mdl созданных моделей.
2. М-файл функции классификации (с комментариями).
3. «Осциллограммы» исходного и синтезированного речевых сигналов.
4. Выводы по работе.

Контрольные вопросы

1. Почему данные о мощности остатка предсказания сначала попадают в блок извлечения квадратного корня и лишь после этого используются в блоке перемножения данных?

2. Как можно объяснить отличие синтезированного сигнала от сигнала на входе речевого кодека?

3. К чему приведет исключение процедуры устранения постоянной составляющей из сигнала возбуждения синтезирующего фильтра?

4. Каким образом осуществляется классификация сегментов речевого сигнала на вокализованные и невокализованные?

Лабораторная работа № 3 Сведение видеoinформации и звука в ролик Возможность использования эффекта «хромакей» Назначение эффекта «морфинг». Сведение видео и звука в ролик. Применение переходных эффектов.

Время выполнения - (8 ч)

Задание 1. Съёмка и монтаж видеофайлов

Самостоятельно выполнить съёмку видеоклипа о себе, Вашей семье, Вашем доме (или общежитии). Видеоклип должен состоять из трёх отдельных сюжетов (частей, сцен). Произвести монтаж (соединение) этих трёх частей. Отснятый материал должен быть «свежим» (нежелательно использовать видеоклипы, снятые до момента выдачи задания на выполнение данной работы).

Задание 2. Монтаж фотографий К видеоклипу, созданному при выполнении первого задания, добавить пять фотографий, импортированных. Тематика фотографий выбирается в соответствии с вариантом и таблицей Таблица 3.2.

Вариант Тема

- 1 Музеи, выставки
- 2 Парки, сады, скверы
- 3 Площади
- 4 Водоёмы
- 5 Театры, концертные залы
- 6 Улицы, проспекты
- 7 Скульптуры
- 8 Растительный мир
- 9 Гостиницы
- 10 Дворцы, дома
- 11 Набережные, пляжи
- 12 Ночной город
- 13 Памятники
- 14 Культовые сооружения
- 15 Площади
- 16 Животный мир

Задание 3. Формирование переходов между клипами В соответствии с вариантом выбрать тип перехода между клипами и реализовать его в фильме, созданном при выполнении

предыдущих заданий.

Таблица 3.3

Варианты	Переходы	Варианты	Переходы	Варианты	Переходы
1	3D	7	Mask	13	Stretch
2	Build	8	Peel	14	Wipe
3	Clock	9	Push	15	3D
4	F/X	10	Roll	16	Build
5	Film	11	Rotate	17	Clock
6	Flashback	12	Slide	18	Mask

Задание 4. Формирование звукового сопровождения В соответствии со своим вариантом выбрать музыкальный файл из папки, местоположение которой указывается преподавателем. Музыкальный файл разместить на монтажном столе (дополнить свой фильм звуковым сопровождением).

Задания 5

1. Разработайте программу, отображающую в окне четыре области с возможностью воспроизведения различных медиа-потоков (видео, звук). Предусмотреть возможность управления воспроизведением каждого потока в отдельности и возможность открытия любых медиа файлов. Исследовать качество воспроизведения файлов различных типов.

2. Создайте приложение, отображающее трехмерный вращающийся куб, на каждой грани которого воспроизводится видео. Реализуйте элементы управления видео (кнопки воспроизведения, остановки, паузы) для каждой грани в отдельности.

3. Создайте приложение, отображающее трехмерный вращающийся куб, на каждой грани которого воспроизводится видео. Реализуйте элементы управления видео (кнопки воспроизведения, остановки, паузы, слайдеры позиции, громкости и скорости воспроизведения) для одной текущей грани. Текущая грань – это грань, повернутая к пользователю и занимающая наибольшую площадь на в окне.

4. Разработайте приложение, собирающее из указанной папки и из всех ее подпапок медиа-контент с возможностью одновременного или последовательного воспроизведения.

5. Разработайте специализированный браузер, собирающей из указанной папки и из всех ее подпапок медиа-контент с возможностью проигрывания указанных элементов. Каждый элемент в браузере отображается в виде изображения. Для реализации этой функции из видео необходимо выбрать один значимый кадр.

6. Разработайте приложение, воспроизводящее последовательно и случайным образом короткие фрагменты видео из различных видео файлов.

7. Реализуйте игру «угадай кадр». Из набора видеофайлов выбирается случайный кадр и игрок должен угадать название фильма. Название фильма содержится в названии файла.

8. Реализуйте тайм линию в медиа-проигрывателе, над которой будут показываться миниатюрные изображения кадров видео для этой позиции там линии.

9. Создайте приложение видеопроигрыватель, с возможностью отбрасывания видеотени позади основного видео. При создании используйте преобразование скос.

10. Создайте приложение с возможностью наложения разных звуковых дорожек с разной громкостью на видео.

11. Разработайте приложение, в котором будет несколько элементов для вывода видео. Все элементы повернуты под разными углами и частично перекрывают друг друга (например, используйте расположение «веером»). Реализуйте функцию проигрывания одного видео с текущей позиции при наведении курсора мышки на этот элемент.

12. Реализуйте игру «составь кадры в хронологическом порядке». Программа выбирает случайным образом кадры из видеопотока. Задача игрока отгадать последовательность воспроизведения кадров в видео.

13. Разработайте программу поиска «ключевых» кадров в видеопотоке. Ключевые кадры это кадры, при которых меняется полностью вся сцена. Для выявления таких кадров ис-

пользуйте признаки резкой смены яркости и контрастности.

14. Создайте приложение, которое отображает ряд кадров из указанного видео и при нажатии на полученный кадр начинается воспроизведение видео с этого кадра.

15. Разработайте приложение, воспроизводящее видео на странице книги. При нажатии на книгу страница переворачивается, а на новой странице воспроизводится видео через одну минуту от текущего. Тем самым реализуется возможность перемотки.

16. Разработайте программу поиска кадров в видеопотоке с какой-либо преобладающей цветовой компонентой. Например, кадров где много синего цвета

Лабораторная работа № 4 Создание титров. Анимация титров. Сохранение готового ролика. Рендеринг (визуализация) ролика. Запись ролика на диск (CDR, CDRW, DVD).

Время выполнения - (14 ч)

Задания к лабораторной работе

1. Разработайте программу, которая при воспроизведении видео, показывает случайным образом области воспроизведения в виде увеличивающихся прямоугольников. Количество одновременно показанных прямоугольников постепенно увеличивается, как и их размер, что приводит постепенно к тому, что видео воспроизводится во все окне.

2. Создайте приложение для отображения видео в окне. Областью воспроизведения является несколько звезд, плавно меняющих свое местоположение и размеры.

3. Реализуйте приложение, в котором пользователь может самостоятельно определить область воспроизведения видео, выбирая и располагая на окне прямоугольники и овалы произвольной величины.

4. Создайте приложение, воспроизводящее видео в окне. Область воспроизведения - круг, плавающий по окну. Реализуйте упругие столкновения с границами окна. Скорость движения и радиус круга может меняться динамически во время воспроизведения.

5. Создайте приложение для отображения видео в окне. Областью воспроизведения является пульсирующее сердце. При воспроизведении плавно меняется размытие.

6. Создайте анимацию на основе пути падения снежинок, листьев и пр. с помощью XAML разметки.

7. Разработайте дизайн приложения с динамически меняющимися элементами. При наведении на элемент может меняться цвет, форма элемента, текстовая надпись, появляться или скрываться тень и т.п.

8. Создайте видеопроигрыватель с вращающейся областью просмотра. Вращение реализуйте через анимацию в XAML разметке.

9. Реализуйте через анимацию в XAML разметке анимацию катящегося мяча по различным наклонным поверхностям. Мяч должен вращаться и двигаться с различной скоростью.

10. Разработайте приложение анимирующее движение видео по ряду кривых Безье внутри окна. Анимация должна реализовываться через XAML разметку.

11. Создайте приложение анимирующее движение вращающегося куба по заданной траектории. При этом используйте только XAML разметку.

12. Разработайте приложение на основе XAML, анимирующее движение корабля по волнам. В сцене должны присутствовать анимированные чайки, облака и солнце.

13. Разработайте программу на основе XAML, анимирующее движение планет в солнечной системе. В сцене должны присутствовать анимированное движение спутников вокруг планет.

14. Реализуйте анимацию батальной сцены с помощью XAML разметки. В сцене должны присутствовать все три типа анимации срабатывающие на различные события.

15. Формирование титров. Для созданного фильма сформировать титры. В титрах указать фамилию, имя, группу студента, дату съемки, тип съёмочной аппаратуры, название видео редактора, название использованного музыкального произведения (звуковое сопровождение). Общая длительность фильма (три клипа, пять фотографий, титры, переходы) должна

быть

Лабораторная работа № 5 Создание презентаций. Использование шаблонов. Элементы навигации. Компоновка и оформление окон сценария. Тестирование презентации. Создание элементов навигации

Время выполнения - (6 ч)

Задание.1 Создать оригинальную презентацию на заданную тему согласно номеру варианта. Сохранить файл как Демонстрацию презентации. Архивировать файл демонстрации в самораспаковывающийся архив со средней степенью сжатия.

Варианты тем презентаций:

1. Стратегическая роль информационных систем в менеджменте
2. Использование информационных систем для бизнес-планирования
3. Реинжиниринг и информационные технологии
4. CASE-технологии и моделирование бизнес-процессов
5. Как информационные технологии способствуют качеству
6. Офис на основе технологии клиент-сервер
7. Системы поддержки работы команд
8. Системы поддержки индивидуальной работы
9. Индивидуальные и групповые модели принятия решений
10. Психологические аспекты работы команд
11. Деловые компьютерные игры в обучении менеджеров
12. Стандарт моделирования бизнес-процессов IDEF
13. Информационный маркетинг
14. Системы рассуждений на основе аналогов
15. Использование статистических пакетов для ответов на вопросы
16. Информационные процессы
17. Информационные системы
18. Информационный менеджмент
19. Управление персоналом в сфере информатизации
20. Информационная среда
21. Организация коммуникативных процессов
22. Информатизация управления персоналом
23. Менеджмент
24. Программа как товар
25. Процесс коммуникаций и эффективность управления
26. Информационная культура
27. Управление капиталовложениями в системе информатизации
28. Системы поддержки принятия решений
29. Виды информационных систем в организациях
30. Менеджеры и системы поддержки управления

Презентация должна состоять из 9-10 слайдов. Поиск информации для создания презентации произвести в Интернете с помощью поисковых машин Rambler, Yandex, Mail, Google и иных. При создании презентации обязательно использовать настройку анимации и иных действий.

Защита лабораторной работы заключается в представлении доклада на заданную тему с использованием презентации. Анализ и оценка работы производится всей группой в ходе дискуссии на основе критериев.

Из вычислительной техники известно, что **триггер** (триггерная система)— класс электронных устройств, обладающих способностью длительно находиться в одном из двух устойчивых состояний и чередовать их под воздействием внешних сигналов(материал Вики-

педии).

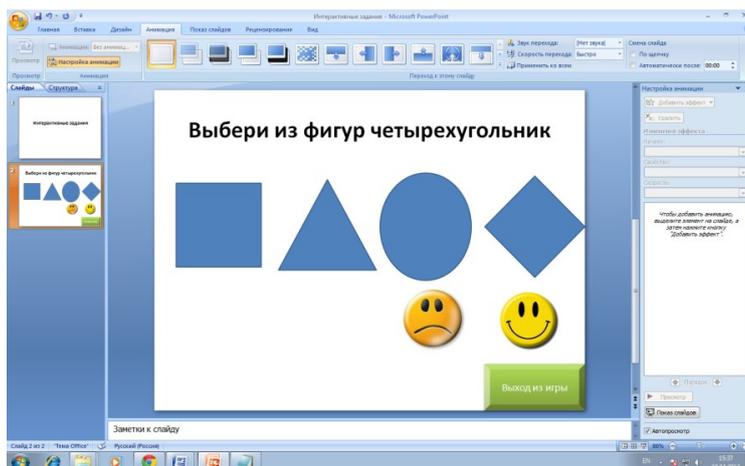
Но, в программе PowerPoint можно создавать различные эффекты анимации. Одним из таких средств является триггер.

Триггер в PowerPoint - средство анимации, позволяющее задать условие действия или времени выделенному элементу. При этом анимация запускается по щелчку.

Задание 2

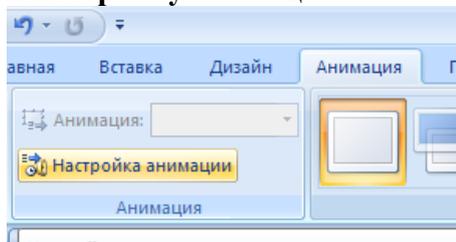
Рассмотрим технологию создания триггера на примере логической игры. Сделаем слайд с заданием: «Выбери из фигур четырехугольник». При щелчке на верную фигуру – фигура увеличивает размер и вылетает веселый смайл, при щелчке на неверную фигуру – вылетает грустный смайл.

1. Вначале нужно на слайде разместить объекты.

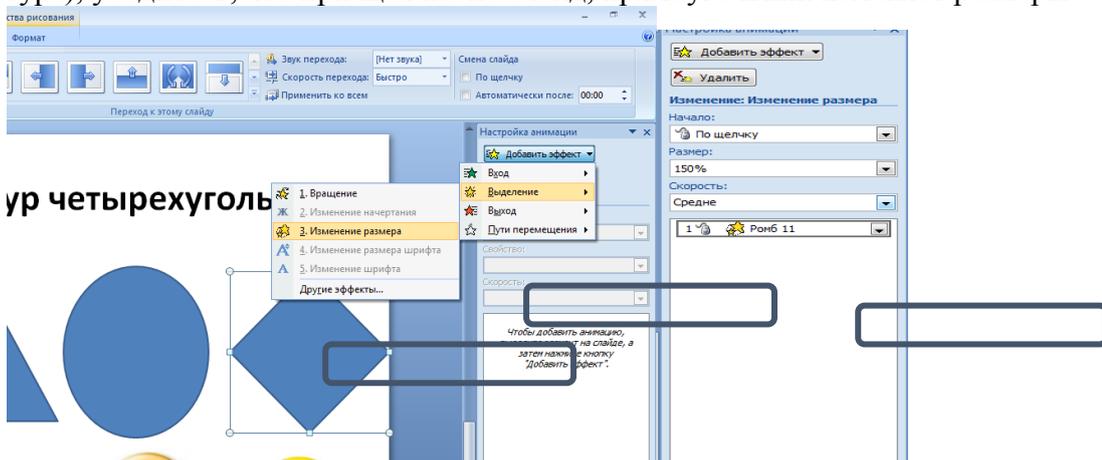


2. Нужно продумать каким образом к ним будет применена анимация и триггер.

Для начала зададим выбранную анимацию. Для этого на верхней панели «Анимация» откроем **настройку анимации**.

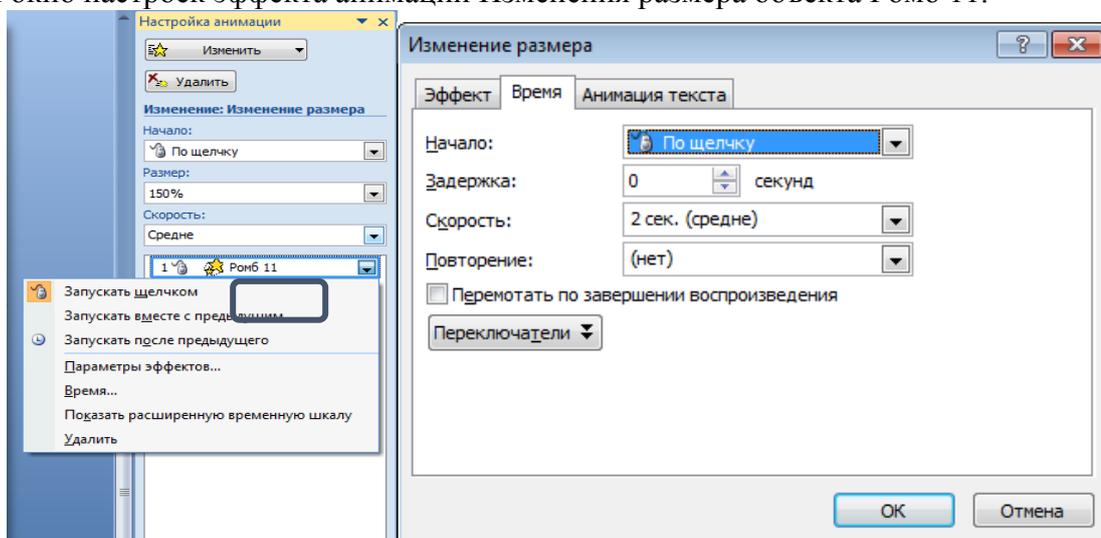


3. Кнопкой мыши выделяем верную фигуру которая будет изменять размер. В области задач (справа), в окне «Настройка анимации» щелкаем по кнопке «Добавить эффект», далее «Выделение» - «Изменение размера». Обязательно эффект анимации должен запускаться щелчком, как на изображении ниже. Просмотрите слайд в режиме просмотра (клавиша F5 на клавиатуре), убедитесь, что при щелчке на слайд, прямоугольник изменяет размеры.

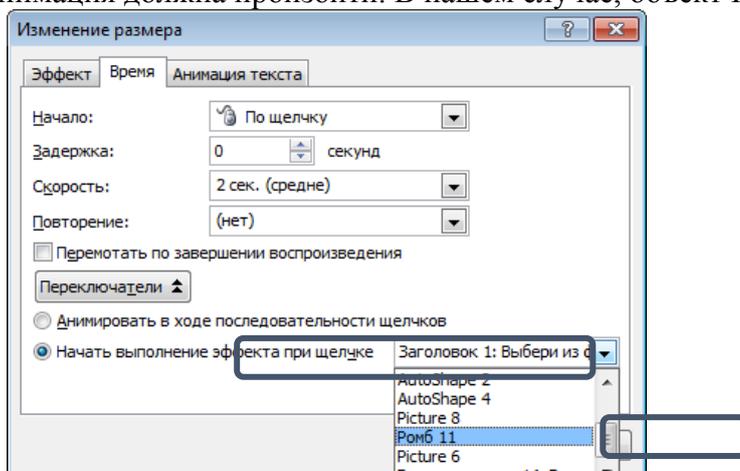


4. Далее настроим так, чтобы при щелчке именно на **прямоугольник** (а не на слайд), он изменял размер, для этого и применяется триггер.

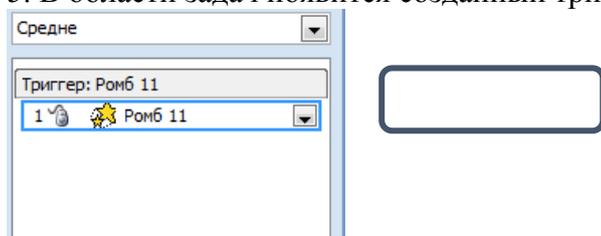
Нажмите на кнопку со стрелкой у эффекта анимации изменение размера (в нашем случае, Ромб 11) на панели **Настройка анимации**. Далее выберите команду **Время...**, появится окно настроек эффекта анимации Изменения размера объекта Ромб 11.



Нажмите кнопку «**Переключатели**» в левой нижней части окна. Выберите параметр «**Начать выполнение эффекта при щелчке**». Будет отображен список объектов на слайде. Выбрать нужный объект из предложенного списка, то есть тот объект, при нажатии на который, анимация должна произойти. В нашем случае, объект Ромб 11.

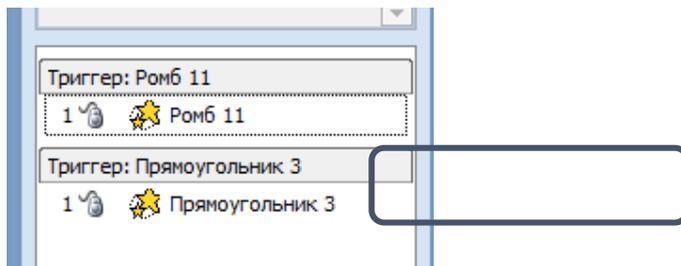


5. В области задач появится созданный триггер на Ромб 11.

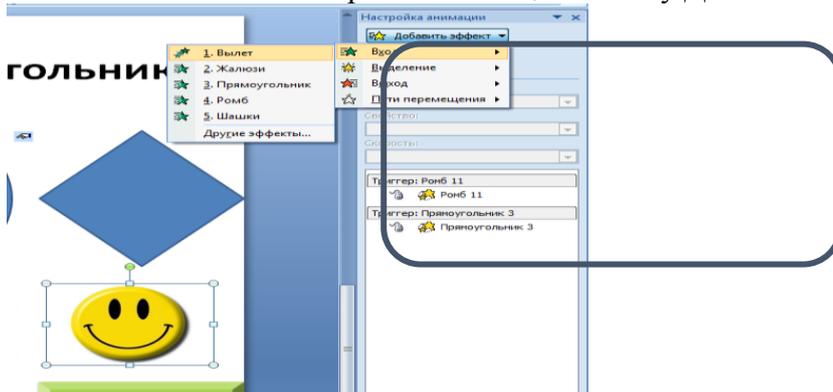


6. Просмотрите слайд в режиме просмотра. При нажатии на прямоугольник – он изменяет размер, но если нажать в другое место, слайд закрывается, чтобы этого не происходило, уберите настройку Смена слайдов по щелчку. Для этого во вкладке Анимация, уберите галочку в блоке Смена слайдов напротив элемента по щелчку. Снова просмотрите слайд в режиме просмотра. Чтобы выйти из режима просмотра, нажмите клавишу Esc.

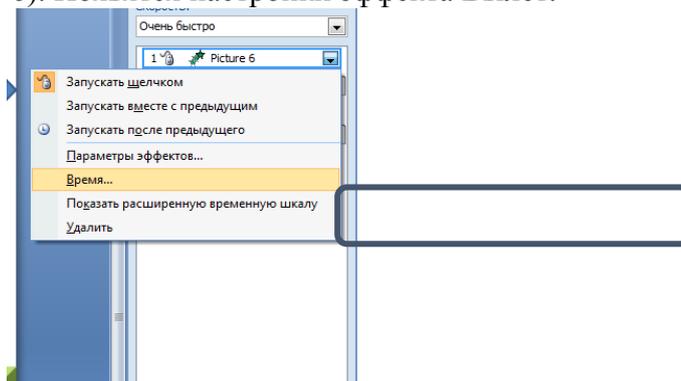
7. Сделайте триггер для второго верного прямоугольника. После этого, у вас на панели задач, должно появиться примерно следующее:



8. Сделаем так, чтобы при нажатии на прямоугольник – он не только изменял размер, но и вылетал веселый смайл. Для этого добавьте эффект анимации для смайла: Выделите смайл, нажмите на панели Настройка анимации кнопку Добавить эффект – Вход-Вылет.

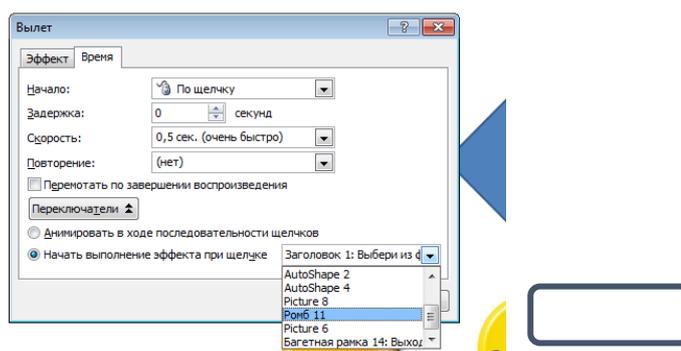


9. Далее нажмите на кнопку со стрелкой у эффекта анимации Вылета (в нашем случае, Picture 6). Появятся настройки эффекта Вылет.

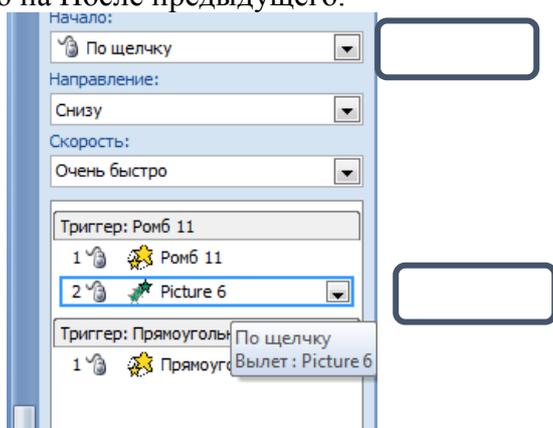


10. Нажмите кнопку «Переключатели» в левой нижней части окна. Выберите параметр «Начать выполнение эффекта при щелчке». Будет отображен список объектов на слайде. Выбрать нужный объект из предложенного списка, то есть тот объект, при нажатии на который, анимация должна произойти. В нашем случае, объект Ромб 11.

из фигур четырехуголь



12. На панели Настройка анимации будет примерно следующее. У одного триггера – два действия. Чтобы второй эффект появлялся вместе с предыдущим, поменяйте настройку Начало на После предыдущего.



13. Посмотрите слайд в режиме просмотра, по щелчку на верный прямоугольник, он должен изменять размер и появляется смайл.

Доделайте задание:

- при щелчке на правильные фигуры, они изменяют размер и появляются смайлы;
- при щелчке на неправильные фигуры появляются грустный смайл;
- при щелчке на кнопку Выход из игры, презентация выходит из режима просмотра

Задание 4

Создайте несколько заданий для своих обучающихся по теме своего учебного предмета с использованием триггеров (для дистанционного обучения). Встройте в эскиз учебной темы работу обучающихся с получившимся ресурсом.

Задание 5

Создать слайд с тестовым заданием, в котором обучающиеся должны выбрать верную букву для ряда слов. При выборе правильно буквы, она встает на место, при выборе неправильной меняет цвет и появляется цифра -1. Чтобы просчитать оценку, ученику нужно вычесть из «пятерки» единицы.

Создайте слайд следующего вида:



Внимание: каждый элемент пишите отдельной текстовой строкой. Выберите шаблон презентации с одним заголовком (или удалите основную текстовую рамку). Выберите на панели инструментов «Надпись» - введите текст.

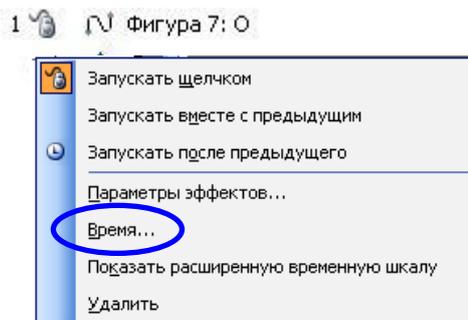
Настройте анимацию на элементы.

Продумайте, как будут появляться или изменяться элементы. Необходимо настроить анимацию: 1) входа для картинок-оценок 2) выделения для верных и неверных ответов.

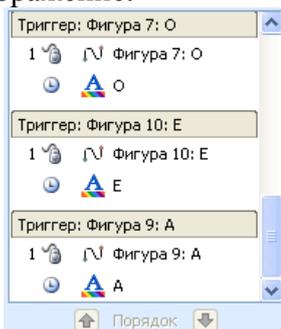
В данном случае, при нажатии на правильный ответ назначена анимация «перемещение влево» и «изменение цвета текста», а при нажатии на неверный ответ – «появление» для картинки "-1" и "контрастирующий цвет" для неверной буквы.

Примените триггер для анимации верных ответов.

Откройте «Настройка анимации» во вкладке Анимация. Щелкните ПКМ на анимацию перемещения буквы О в первом вопросе. Выберите вкладку «Время».

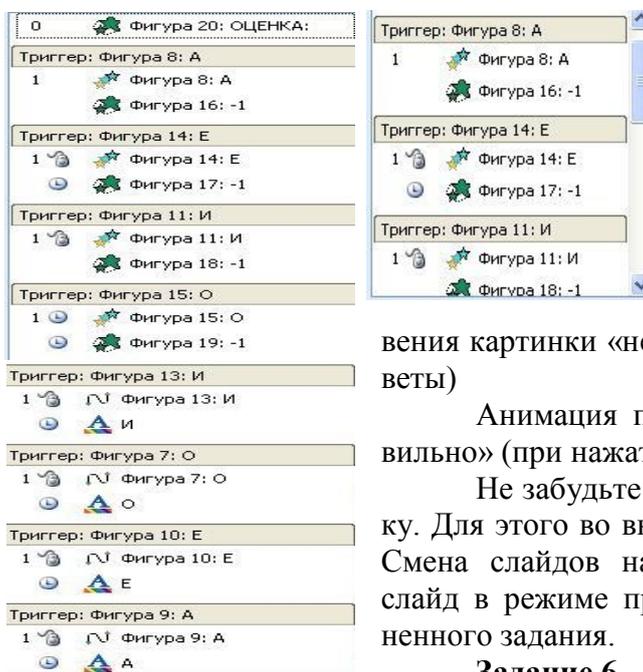


В появившемся окне выберите кнопку «Переключатели» - «начать выполнение эффекта при щелчке». Выберите элемент слайда, при нажатии на который применится анимация (в нашем случае О). В поле «Настройка анимации» появится запись «триггер», а на слайде изображение:



Аналогично назначьте триггер на все правильные ответы.

Примените триггер для неправильных ответов. В нашем случае при нажатии на неправильный ответ должно произойти изменение цвета и появление «-1». Выделите обе анимации (удерживая Ctrl нажмите на одну и вторую анимации в списке), далее по шагу 3. При



правильном выполнении для каждой неверной буквы ответа создается триггер.

Вид «неправильных ответов» на панели «Настройка анимации»:

Вид окна «Настройка анимации» после применения всех триггеров:

Анимация появления слова «оценка» (может отсутствовать)

Анимация появления и исчезновения картинки «неправильно» (при нажатии на неверные ответы)

Анимация появления и исчезновения картинки «правильно» (при нажатии на верный ответ)

Не забудьте убрать настройку Смена слайдов по щелчку. Для этого во вкладке Анимация, уберите галочку в блоке Смена слайдов напротив элемента по щелчку. Запустите слайд в режиме просмотра, проверьте правильность выполненного задания.

Задание 6

Создайте несколько тестовых заданий (тренажеров) по теме своего учебного предмета. Подумайте, в каких условиях, вы могли бы применить данный учебный материал в образовательном процессе. Встройте в эскиз учебной темы работу обучающихся с получившимся ресурсом.

Контрольные вопросы.

1. Понятие презентации и слайда
2. Состав слайда
3. Классификация презентаций

4. Этапы создания презентации
5. Презентационные программы
6. Способы создания презентации
7. Рекомендации по созданию презентации
8. Задачи, решаемые опубликованием презентации
9. Что представляет собой компьютерная презентация?
10. Какими способами можно создавать презентации?
11. Какие функциональные возможности использует Microsoft Power Point для оформления созданной презентации?
12. Какие гиперссылки используются в презентации?
13. На какие объекты можно делать гиперссылки?
14. В каких презентациях используются гиперссылки?
15. Что такое шаблон оформления?
16. Зачем необходима функция скрытый слайд?
17. Какие есть виды форматирования слайдов?

3. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

Самостоятельная работа студентов является обязательной для каждого студента, а её объём определяется учебным планом. Формы самостоятельной работы студентов определяются содержанием учебной дисциплины, степенью подготовленности студентов.

Самостоятельная работа — одна из важнейших форм овладения знаниями. Самостоятельная работа включает многие виды активной умственной деятельности студента: слушание лекций и осмысленное их конспектирование, глубокое изучение источников и литературы, консультации у преподавателя, написание реферата, подготовка к практическим и лабораторным занятиям, экзаменам, самоконтроль приобретаемых знаний и т.д.

Изучение дисциплины следует начинать с проработки настоящей рабочей программы, особое внимание уделяя целям и задачам, структуре и содержанию курса.

Студентам рекомендуется получить в научной библиотеке университета учебную литературу по дисциплине, необходимую для эффективной работы на всех видах аудиторных занятий, а также для самостоятельной работы по изучению дисциплины.

Теоретический материал курса становится более понятным, когда дополнительно к прослушиванию лекции и изучению конспекта, изучаются и книги. Легче освоить курс, придерживаясь одного учебника и конспекта. Рекомендуется, кроме «заучивания» материала, добиться состояния понимания изучаемой темы дисциплины. С этой целью рекомендуется после изучения очередного параграфа выполнить несколько простых упражнений на данную тему. Кроме того, очень полезно мысленно задать себе следующие вопросы (и попробовать ответить на них): о чем этот параграф, какие новые понятия введены, каков их смысл, что даст это на практике?

При организации самостоятельной работы студенты особое внимание должно уделяться анализу учебно-методической литературы по дисциплине.

Процесс организации самостоятельной работы студентов включает в себя следующие этапы:

1. Подготовительный (определение целей, составление программы, подготовка методического обеспечения, подготовка оборудования).
2. Основной (реализация программы, использование приемов поиска информации, усвоения, переработки, применения, передачи знаний, фиксирование результатов, самоорганизация процесса работы).
3. Заключительный (оценка значимости и анализ результатов, их систематизация, оценка программы и приемов работы, выводы о направлениях оптимизации труда).

Тематика заданий для самостоятельной работы представлена в рабочей программе.