

У даній роботі проводиться дослідження сучасних цифрових аудіо-технологій, що використовуються при розробці мультимедійних видань. На основі проведеного дослідження здійснюється оцінка впливу звуку на органи слуху людини та аналіз сприйняття нею синтетичної мови

Ключові слова: електронне видання, мультимедіа, звук, синтез мови

В данной работе проводится исследование современных цифровых аудиотехнологий и эффективности их использования при разработке мультимедийных изданий. На основе проведенного исследования осуществляется оценка влияния звука на органы слуха человека и анализ восприятия им синтетической речи

Ключевые слова: электронное издание, мультимедиа, звук, синтез речи

ИССЛЕДОВАНИЕ ЗВУКА КАК МУЛЬТИМЕДИЙНОГО КОМПОНЕНТА ЭЛЕКТРОННЫХ ИЗДАНИЙ

И. Н. Егорова

Кандидат технических наук, доцент*

Контактный тел.: (057) 702-13-78

E-mail: irinaiegorova@gmail.com

М. М. Чиркина*

Контактный тел.: 096-623-09-42

E-mail: chirkinamargarita@rambler.ru

*Кафедра мультимедийных систем и технологий

Харьковский национальный университет радиоэлектроники

пр. Ленина, 14, г. Харьков, Украина, 61166

1. Введение

Современная система образования все активнее использует информационные технологии и компьютерные телекоммуникации. Особенно динамично развивается система дистанционного образования как наиболее востребованная и современная. Основным носителем информации при дистанционной форме обучения выступают электронные издания. Их достоинствами являются мобильность, доступность в связи с развитием компьютерных сетей, адекватность уровню развития современных научных знаний. Онлайн технологии, широко используемые электронными изданиями, дают возможность пользователю самостоятельно выбирать удобное для него время и место для занятий. Информация, содержащаяся в электронных изданиях, требует постоянного обновления. Кроме того, при помощи электронных изданий осуществляется контроль знаний - компьютерное тестирование.

Дистанционное образование базируется на использовании мультимедийных технологий. Внедрение мультимедиа-технологий в образовательный процесс является одним из ключевых моментов информатизации образования. В настоящее время мультимедиа-технологии динамично развиваются и относятся к одному из наиболее перспективных направлений информационных технологий [1]. Актуальность их применения обусловлена информатизацией общества, широким распространением сети Интернет, а также стремительным развитием дистанционного образования.

2. Анализ проблемы

Мультимедиа – (от англ. multi - много и media – среда) компьютерная технология, обеспечивающая

возможность создания, хранения и воспроизведения разнородной информации, включая текст, звук и графику (в том числе движущееся изображение и анимацию).

Мультимедийные издания в настоящее время стали неотъемлемой частью жизни многих учащихся. Они зачастую воспринимают эти издания с гораздо большим интересом, чем обычный школьный учебник. Издания, разработанные с использованием мультимедиа-технологий, позволяют в значительной степени повысить уровень усвоения материала учащимися. Интерактивность является очень важной составляющей мультимедиа. Согласно статистическим исследованиям, люди запоминают только 20% того, что они видят, и 30% того, что они слышат. Известно также, что уровень восприятия видео- и аудио-составляющих - 50%, и целых 80% - уровень восприятия видео-, аудио- и интерактивной составляющих одновременно (рис. 1).

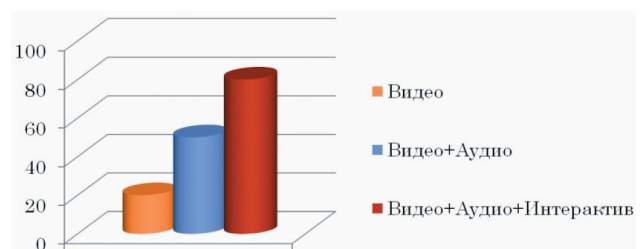


Рис. 1. Уровень восприятия мультимедийной информации человеком

Очевидно, что звуковое сопровождение играет значительную роль в процессе восприятия информации. Простейший способ вывода звукового сообщения заключается в записи речевого сигнала в

аналоговой форме на магнитную ленту и его воспроизведению по управляющему сигналу. Однако, это дорогостоящий и трудозатратный способ. Значительно более удобны цифровые системы, в которых речевой сигнал записывается в дискретной форме [2]. При этом имеется в виду прочтение вслух произвольной текстовой информации – синтез речи, а не проигрывание предварительно записанных звуковых файлов.

Использование таких систем требует решения ряда проблем, таких как:

- автоматический синтез речи по тексту;
- воспроизведение характеристик человеческого голоса;
- анализ восприятия человеком синтетической речи.

3. Цель и задачи исследования

Целью работы является исследование цифровых технологий, используемых при создании звукового сообщения. Для этого необходимо исследовать методы синтеза речи, а также провести анализ влияния звука на органы слуха человека и восприятия им синтетической речи.

4. Исследование методов синтеза речи

Под синтезом речи понимают восстановление формы речевого сигнала по его параметрам. В работе исследуются основные методы синтеза речи, такие как: параметрический, конкатенативный (компиляционный) и синтез по правилам [3].

В первом случае речевой сигнал представляется набором небольшого числа непрерывно изменяющихся параметров. Этот тип синтеза целесообразно применять в тех случаях, когда набор сообщений ограничен и изменяется не слишком часто. Качество параметрического синтеза может быть очень высоким (в зависимости от степени сжатия информации в параметрическом представлении). Достоинством такого способа является возможность записи речи на любом языке любым диктором. Однако параметрический синтез не может применяться для произвольных, заранее не заданных сообщений.

Во втором случае синтез сводится к составлению сообщения из предварительно записанного словаря исходных элементов. Очевидно, что содержание синтезируемых сообщений ограничивается объемом словаря.

Основная проблема в компилятивном синтезе – объемы памяти для хранения словаря. В этой связи используются разнообразные методы сжатия/кодирования речевого сигнала.

И, наконец, синтез по правилам обеспечивает управление всеми параметрами речевого сигнала и, таким образом, может генерировать речь по заранее неизвестному тексту. В этом случае параметры, полученные при анализе речевого сигнала, сохраняются в памяти так же, как и правила соединения звуков в слова и фразы. Синтез реализуется путём моделирования речевого тракта, применения анало-

говой или цифровой техники. Этот метод базируется на запрограммированном знании акустических и лингвистических ограничений и не использует непосредственно элементы человеческой речи.

В системах, основанных на способе синтеза по правилам, выделяется два подхода, один из которых направлен на построение модели речеобразующей системы человека и известен как артикуляторный синтез, а второй – формантный синтез по правилам.

В настоящее время наилучшее качество речи в классе систем синтеза речи по правилам обеспечивают формантные синтезаторы. Они исследованы лучше других. Анализ результатов этих исследований дает важную информацию о возможностях, недостатках и направлениях дальнейших исследований в области синтеза речи.

Разборчивость и натуральность формантного синтеза речи достигли достаточно высокого уровня, приемлемого в условиях практического применения, однако установлено, что даже для формантных синтезаторов с наиболее высокой разборчивостью, восприятие синтетической речи требует больше умственных усилий и внимания, чем восприятие естественной речи [4].

5. Анализ восприятия человеком синтетической речи

Несмотря на кажущееся качество произношения, воспроизводимая синтезаторами речь воспринимается человеком с трудом. То есть достаточно качественный синтезатор с хорошо подобранной фонетической базой будет прекрасно восприниматься на слух в течение 15 – 20 минут, но потом абсолютное большинство людей перестает воспринимать смысл того, что произносится.

А происходит это из-за того, что для прослушивания синтезированной речи используются дополнительные центры обработки головного мозга, и он просто устает. Таким образом, головной мозг не воспринимает синтезированную речь как естественную, которая сразу обрабатывается в речевом центре.

Вторая проблема – личное восприятие произносимого текста. Когда текст читает человек, он поневоле пропускает смысл воспроизводимого через себя, и слушатели чувствуют его отношение в интонациях и нюансах.

Современные программы не обладают такой возможностью, хотя самые передовые из них пытаются имитировать интонацию путем модуляции тембра, длительности фонов и пауз. Но мозг быстро устает исправлять огрехи воспроизведения, и слушатель теряет нить повествования.

Очевидно, для решения этой задачи требуются технологии из области создания искусственного интеллекта, а именно – «извлечение смысла» из воспроизводимого текста.

Третья проблема – низкая помехоустойчивость синтезированной речи. Как показывают эксперименты достаточно лишь небольшого шума, чтобы слушатель перестал воспринимать смысл произносимого синтезатором. Это объясняется тоже достаточно просто. Так как для обработки синтезированной речи головной мозг использует дополнительные центры, то

при наличии постороннего шума, разговора или необходимости делать какую-то работу, мозг слушателя просто не справляется («перегружается») и человек перестает понимать смысл произносимого. Эффект помех существенно ограничивает возможности применения синтезатора в реальных условиях техногенных, производственных и других шумов.

6. Выводы

В данной работе проводится исследование современных цифровых аудиотехнологий и эффективности их использования при разработке мультимедийных изданий. Рассмотрены наиболее распространенные программные системы, такие как *Govorilka*, *Sakrament Talker*, *Balabolka*, *ICE Book Reader Pro*, *Cool Reader*,

Рассказчик. Все названные приложения используют метод преобразования TTS (Text to Speech), позволяющий представлять фрагменты текста в виде разговорной речи.

На основе проведенного исследования осуществляется оценка влияния звука на органы слуха человека и анализ восприятия им синтетической речи. Наиболее эффективным является применение анализируемого метода для озвучивания словарей и справочных сообщений.

Проведенное исследование позволяет сделать вывод, что звуковые сообщения, созданные с помощью синтезаторов речи должны быть небольшими по длительности и использоваться в качестве вспомогательных комментариев, а не озвучивать весь объем информации. Полученные результаты используются при разработке электронного учебного пособия.

Литература

1. Ясницкий, Ю. А. Мультимедиа-технологии – новый уровень качества профессионального образования [электронный ресурс]. – Режим доступа: [www URL: http://festival.1september.ru/articles/511345/index.html](http://festival.1september.ru/articles/511345/index.html) - 2012г.
2. Савчук, В. Л. Электронные средства сбора, обработки и отображения информации [электронный ресурс]. – Режим доступа: [www URL: http://www.ie.tusur.ru/books/COI/index.html](http://www.ie.tusur.ru/books/COI/index.html).
3. Фролов А., Фролов Г. Синтез и распознавание речи. Современные решения [электронный ресурс]. – Режим доступа: [www URL: http://www.frolov-lib.ru/books/hi/index.html](http://www.frolov-lib.ru/books/hi/index.html) - 2003г.
4. Сорокин, В. Н. Синтез речи [текст]. – М.: Наука, Гл. ред. физ-мат. лит., 1992. – 392 с.

Abstract

Introduction of multimedia technologies in educational processes is one of key points of informational education. Currently, multimedia technologies are developing dynamically and belong to one of the most promising areas of informational technologies. Urgency of multimedia technologies application is determined due to computerization of society, wide spreading of the Internet and fast development of distance learning.

In this article we have examined digital technologies of multimedia creation, investigated methods of speech synthesis, analyzed influence of sound onto human's hearing organs, and analyzed human's perception of synthetic speech. Besides the most common software systems, such as Govorilka, Sakrament Talker, Balabolka, ICE Book Reader Pro, Cool Reader, Rasskazchik were examined. All of these applications use a method of converting TTS (Text to Speech), which represents fragments of the text in the form of speech.

The most effective is the application of the analyzed method for voicing dictionaries and reference messages. The obtained results are used in the development of electronic editions

Keywords: electronic edition, multimedia, sound, speech synthesis