

КОМПЬЮТЕРНАЯ СИСТЕМА ВИРТУАЛЬНОГО ОБЩЕНИЯ ЛЮДЕЙ С ПРОБЛЕМАМИ СЛУХА

Юрий Крак, Александр Бармак, Александр Ганджа,
Антон Тернов, Николай Шатковский

Аннотация: В статье представлена комплексная информационная технология для бессловесной коммуникации между людьми с проблемами слуха, основанная на жестомимическом языке.

Ключевые слова: симуляция, язык жестов, компьютерная система.

ACM Classification Keywords: I.2.8 Problem Solving, Control Methods, and Search H.1.1 Systems and Information.

Conference: The paper is selected from XIVth International Conference "Knowledge-Dialogue-Solution" KDS 2008, Varna, Bulgaria, June-July 2008

Введение

В мире количество полностью глухих людей и людей с тяжелыми формами потери слуха составляет около 1,5% от общей численности населения – это десятки миллионов людей, для которых необходимо создавать средства равноценного общения в обществе. Пункт 7 Правила 5 Приложения к Резолюции ООН 48/96 "Стандартные правила обеспечения равных возможностей для инвалидов" гласит: "Необходимо обеспечить, чтобы язык жестов применялся для обучения глухих детей, в их семьях и сообществах. Необходимо также обеспечить услуги сурдоперевода для того, чтоб способствовать общению глухих с другими людьми".

В Украине больше полумиллиона детей с недостатками слуха, а количество глухих людей, для которых необходимо разрабатывать современные средства обучения и общения соответственно к мировому научно-техническому развитию – составляет миллионы. Развитие современной науки, компьютеризация общества, использование мультимедийных и Интернет технологий создали достаточные условия для разработки компьютерных систем коммуникации этих людей в формах и образах близких и понятных для них и для окружающего мира.

Основной формой общения глухих есть язык жестов. Язык жестов имеет национальные особенности (например, английский, французский, украинский, русский и другие языки), причем люди с недостатками слуха используют при общении два жестовых языка, которые имеют разную грамматику и разный набор жестов:

- разговорный язык жестов, который используется в повседневном общении и имеет собственную грамматику, достаточно отличающуюся от естественного разговорного языка;
- калькулирующий язык жестов, который используется в официальной и деловой обстановке и содержит в себе как знаки разговорного языка жестов, так и знаки дактильной азбуки, которая служит для отображения слов по буквам, причем калькулирующий язык жестов не имеет своей грамматики, он подчиняется грамматическим правилам национального языка.

Важной проблемой общения глухих с другими людьми есть умение распознать разговорный язык по губам. С этой точки зрения задачи визуального синтеза языка и распознавания по губам есть альтернативой языкового общения для людей с проблемами слуха. Кроме этого, развитие направления автоматического чтения по губам поможет улучшить показатели существующих систем распознавания

языка благодаря получению дополнительного независимого канала информации, а синтезированное озвучивание текстов позволит включить в коммуникационный процесс и людей с проблемами зрения.

В статье рассматривается комплексная информационная технология невербального общения людей с проблемами слуха, как между собой, так и с другими людьми.

Информационная технология невербального общения людей с недостатками слуха

После проведенного анализа структуры языка жестов (на основе анализа американского жестового языка (ASL) [Stokoe, 1960, Stokoe, 2001]), современного состояния проблемы общения людей с ограниченным слухом, была разработана концепция комплексной информационной технологии.

Комплексная информационная технология включает реализацию следующих возможностей:

- модуль перевода обычного текста на язык жестов глухих, который содержит: анимацию процесса проговаривания разговорного и калькулирующего языка жестов с использованием виртуальных трехмерных моделей людей;
- анимацию мимики лица модели (с учетом эмоциональных составляющих) при проговаривании;
- озвучивание (синтез) обычного текста в его звуковой аналог (с использованием разных голосов, с реализацией функций «громче/тише», «дальше/ближе»);
- модуль распознавания по изменению мимики губ текста, который проговаривается.

Концепция экспериментальной технологии виртуального общения людей с проблемами слуха представлена на схеме (рис. 1).



Рис. 1. Схема концепции экспериментальной технологии виртуального общения глухих людей

В рамках этих исследований была реализована модель синтеза украинского языка жестов с использованием технологии Microsoft Agent. Был создан агент, который продуцирует язык жестов, используя некоторое множество жестов. Полученный агент может быть интегрирован как в офисные, так и в Web-приложения (Рис. 2).



Рис. 2. Демонстрація моделювання мови жестів з допомогою технології Microsoft Agent

С использованием определенного множества жестов было реализовано приложение для мобильного устройства (рис. 3)



Рис.3. Приложение для мобильного устройства

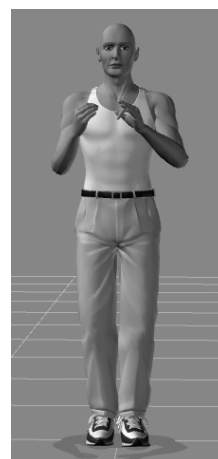


Рис.4. Трехмерная модель генерирующая украинский язык жестов

Для синтеза трехмерной анимации языка жестов построены геометрические классы векторов-образов жестов. Построение этих классов базировалось на использовании технологии motion capture [Menache, 2000]. Motion capture – это технология представления движений, позволяющая перейти от характеристик движения в реальном мире, к фиксации изменений параметров для математической модели. Для автоматического получения необходимых ключевых координат жеста использовалась технология трекинга [Avidan, 2001].

Для представления жеста использовался формат BVH, с последующим экспортом полученных данных на скелетную трехмерную модель (например, в модуле Character Studio для 3D studio MAX или в Poser) .

Предложенная реализация технологии motion capture для фиксации движений жестового языка включает в себя (рис. 5):

1. Получение видео-потока жестикуляции в двух ракурсах: фронт, профиль (все необходимые параметры съемки известны).
2. Обработка видео-потока с выделением следов и положений (координат) рук жестикуляции в пространстве.
3. На основе полученных координат жестикуляции - формирование BVH-файла для синтеза трехмерной анимации.
4. Применение BVH-файла для создания анимации (в Character Studio для 3D studio MAX или в Poser).

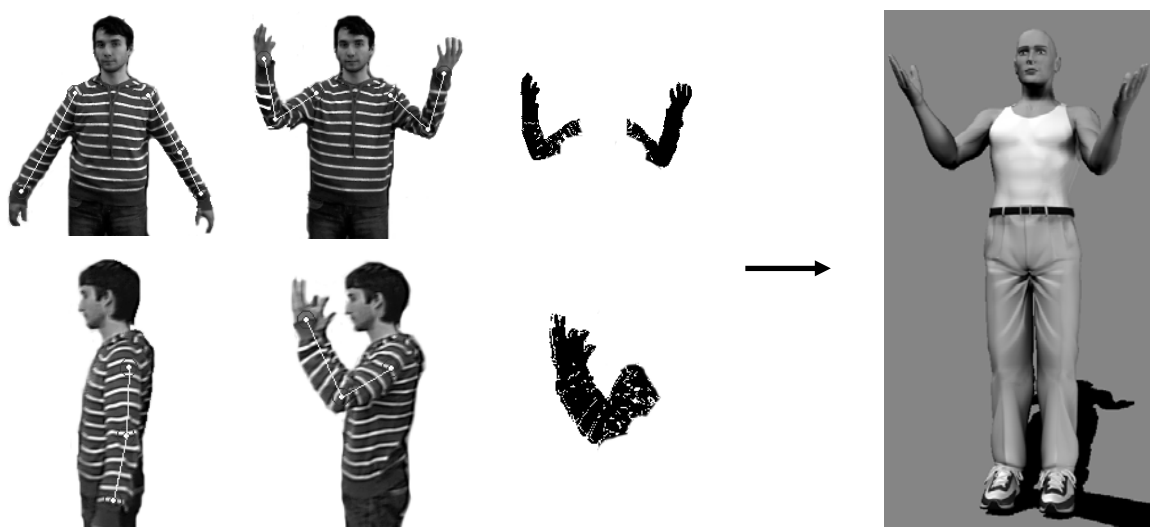


Рис. 5. Построение анимации жеста

Для обработки входного текста (расстановки ударений, выделения инфинитивов, поиска синонимов и типовых фраз языка) разработана информационная модель украинского языка. Модель представлена в виде таблиц реляционной базы данных с набором сохраненных процедур, которые реализуют необходимую, для рассматриваемой технологии, функциональность (рис. 6).

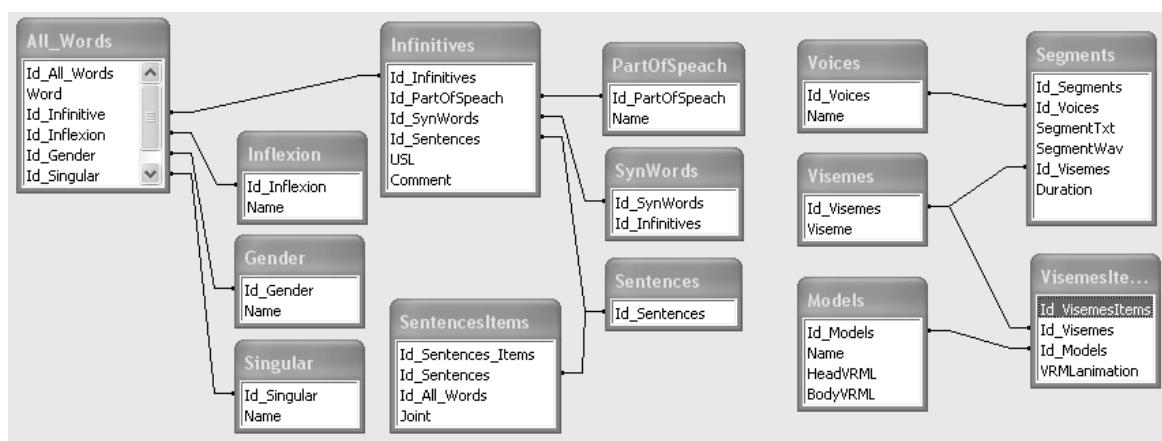


Рис. 6. Логическая схема базы данных – информационная модель украинского языка.

Информационная модель украинского языка содержит более двух с половиной миллионов слов (все возможные словоформы), ударения, синонимы, идиомы, фрагменты слов для синтеза разными голосами, вектора-образы жестов.

Для реализации функции визуализации озвучивания текста разработан синтезатор украинского языка. Синтезатор позволяет создавать голосовой аналог произвольного текста различными голосами с возможностью управления характеристиками голоса (громкость, дальше/ближе). Синтезатор также дает возможность построить визуальное представление процесса проговаривания (как с помощью двумерных визем, так и на трехмерной модели) (рис. 7,8).

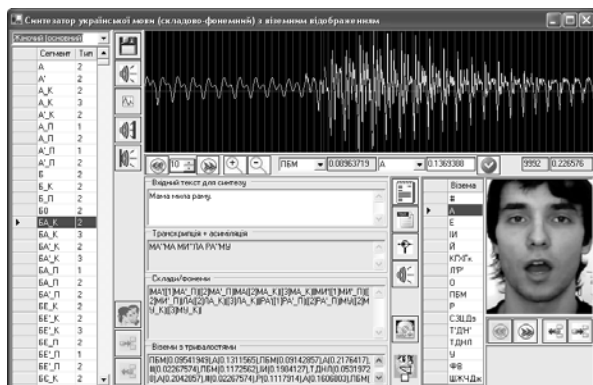


Рис. 7. Синтезатор українського мови



Рис. 8. Синтез визем українського мови

Выводы

Дальнейшие исследования направлены на реализацию всего множества жестов украинского языка жестов с возможностью визуализации на трехмерных моделях (рис. 4). Создается информационная технология и соответствующее программное обеспечение для реализации приведенной концепции компьютерной системы виртуального общения людей с проблемами слуха.

Библиография

- [Stokoe, 1960] Stokoe, W. Sign language structure: An outline of the visual communication systems of the American Deaf. Studies in linguistics, occasional papers 8. Silver Spring, MD: Linstok Press., 1960. - 94 p.
- [Stokoe, 2001] Stokoe W. C. Language in Hand: Why Sign Came Before Speech. Washington, DC: Gallaudet Univ.Press., 2001. - 227p.
- [Menache, 2000] Menache A., Understanding Motion Capture for Computer Animation and Video Games, Morgan Kaufmann, 2000. - 238 p.
- [Avidan, 2001] S. Avidan, "Support vector tracking," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Kauai, Hawaii, volume I, 2001. - P. 184–191.

Authors' Information

Yuriy Krak – The Institute of Cybernetics of National Academy of Science of the Ukraine, the senior scientist, address: 40 Glushkov ave., Kiev, Ukraine, 03680; e-mail: krak@unicyb.kiev.ua

Alexander Barmak – The Institute of Cybernetics of National Academy of Science of the Ukraine, the senior scientist, address: 40 Glushkov ave., Kiev, Ukraine, 03680; e-mail: barmak@svitonline.com

Alexander Gangha – The Institute of Cybernetics of National Academy of Science of the Ukraine, the engineer, address: 40 Glushkov ave., Kiev, Ukraine, 03680

Anton Ternov – The Institute of Cybernetics of National Academy of Science of the Ukraine, the junior scientist, address: 40 Glushkov ave., Kiev, Ukraine, 03680

Nikolai Shatkovskii – The Institute of Cybernetics of National Academy of Science of the Ukraine, the junior scientist, address: 40 Glushkov ave., Kiev, Ukraine, 03680