

О РАСПОЗНАВАНИИ ЖЕСТОВ ЯЗЫКА ГЛУХИХ

Александр Воскресенский (avosj@yandex.ru), Москва
Сергей Ильин (mail@mocarprus.ru),
«Академия фантазий», Москва
Milos Zelezny (zelezny@kky.zcu.cz),
University of West Bohemia, Plzeň, Czech Republic

*Данное исследование проводится при поддержке Министерства Образования, Молодёжи и
Спорта Чешской Республики в рамках совместного проекта DIMAS-CZ, No. ME08106*

Цель работы

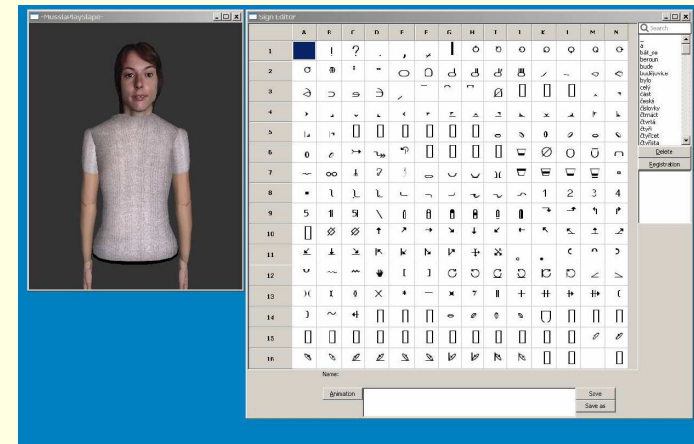
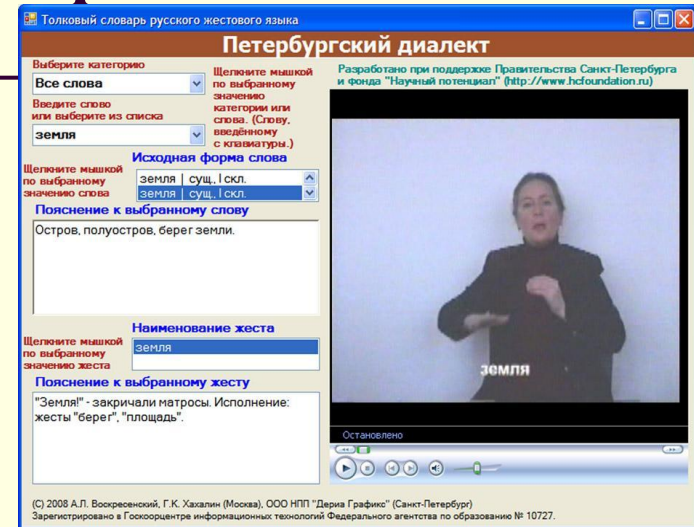
- Положительная социализация глухих людей в условиях современного модернизируемого и инновационного общества
- Облегчение коммуникации между глухими и слышащими, содействие реализации положений Конвенции о правах инвалидов, принятой Генеральной Ассамблеей ООН 13 декабря 2006 года (к которой Россия присоединилась 24 сентября 2009 года)

Средства достижения:

- Описание жестового языка (словари, грамматика)
- Создание обучающих пособий по жестовому языку
- Создание и использование технических средств для демонстрации и восприятия жестов

Предшествующие разработки

- Мультимедийный толковый словарь русского жестового языка RuSLED (Россия)
- 3D Редактор жестов SignEditor с использованием анимированного персонажа, управляемого с помощью нотации HamNoSys (Чехия)



Недостатки предыдущих работ

- RuSLED: в значительной мере содержит жесты «кальки» (Signed Russian), отличные от используемых в естественной жестовой речи (Russian Sign Language); видеофрагменты требуют много памяти и недостаточно информативны
- SignEditor: трудоемкий процесс наполнения словаря, не всегда обеспечивается правильное представление жеста

Пути преодоления: использование метода захвата движений (motion capture) для наполнения словаря 3-мерными изображениями жестов

Примеры нотации



форма кисти	направление кисти	расположение кисти	движение рук
<p><i>Адрес ролика:</i> http://www.sign-lang.uni-hamburg.de/projekte/hamnosys/hamnosyserklaerungen/Bilder/muehsam.mov</p>			

Общая структура представления жеста в HamNoSys

(на примере немецкого жеста для понятия «трудный»)

Запись в нотации HamNoSys чешского жеста «семьдесят»

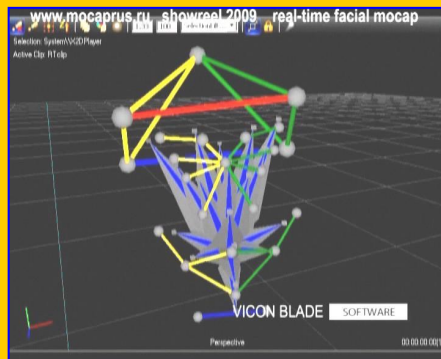
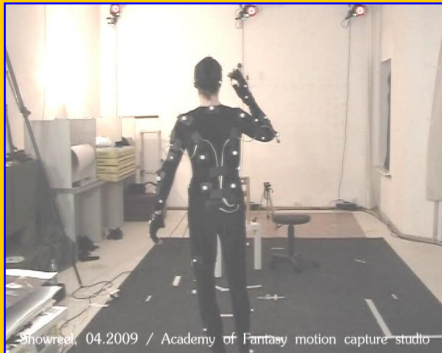
Příklad: znak „sedmdesát“

[] (= []

Пример нотации Л.С. Димскис

лето — П < ↓, папа — П < ↓,
 где П — правая рука, — форма руки.

Съёмка жестов методом захвата движений (motion capture)



Зачем нужен автоматизированный сурдоперевод?

- Недостаточное владение глухими словесным русским языком. Из-за ограниченного активного словаря они часто плохо понимают словесную и письменную речь. Из-за затруднений коммуникации глухие, например, *НИКОГДА* не обращаются к психологам и психиатрам. В свою очередь, слышащие практически не знакомы с жестовым языком.
- Нехватка опытных сурдопереводчиков, владеющих не только «калькой», но и «естественным» жестовым языком
- Требование глухих о признании их лингвистических прав и придании жестовому языку официального статуса

Перевод жестового высказывания

- Выделение ключевых кадров из видеоизображения жестового выражения
- Преобразование растрового изображения кадра в совокупность векторных объектов
- Анализ найденных объектов для определения значений местоположения и конфигурации рук, характеристик мимики и артикуляции
- Разбиение жестового выражения на отдельные жесты, определение грамматической формы и значения каждого жеста
- Построение текста (словесной фразы или фраз), передающего значение жестового высказывания

Распознавание – поиск в пространстве признаков

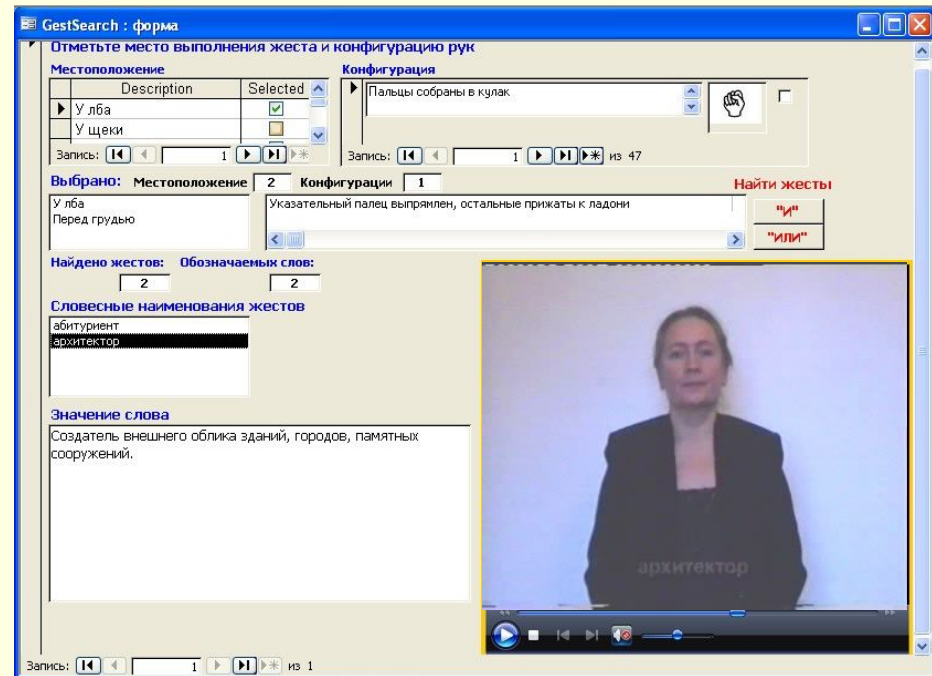
- Сокращение размерности: растр → векторные объекты → кортежи признаков, характеризующих объекты
- Жест – упорядоченное множество кортежей признаков
- Поиск жеста в библиотеке (словаре):
 - Разбиение жестовой фразы на составляющие жесты
 - Учет индивидуальных особенностей демонстрации жестов (допустимые отклонения от эталона)

Поиск жестов в словаре RuSLED (опция «И»)

$$f_1: L \rightarrow G; f_2: K \rightarrow G$$
$$G_f = G_1 \cap G_2,$$
$$G_1 = f_1(\{x: x_1 \text{ } \textcircled{D} \text{ } x_2 \text{ } \textcircled{D} \text{ } \dots \text{ } \textcircled{D} \text{ } x_l \text{ } \dot{Y} \text{ } L_w\});$$
$$G_2 = f_2(\{x: x_1 \text{ } \textcircled{D} \text{ } x_2 \text{ } \textcircled{D} \text{ } \dots \text{ } \textcircled{D} \text{ } x_k \text{ } \dot{Y} \text{ } K_w\}).$$

Выбираются жесты,
содержащие все заданные
позиции и конфигурации.

Если число заданных значений
превышают число позиций или
конфигураций для какого-либо
жеста, то этот жест не будет
включен в множество
результатов G_f , даже если он
содержит указанные
пользователем атрибуты.



Поиск жестов в словаре RuSLED (опция «ИЛИ»)

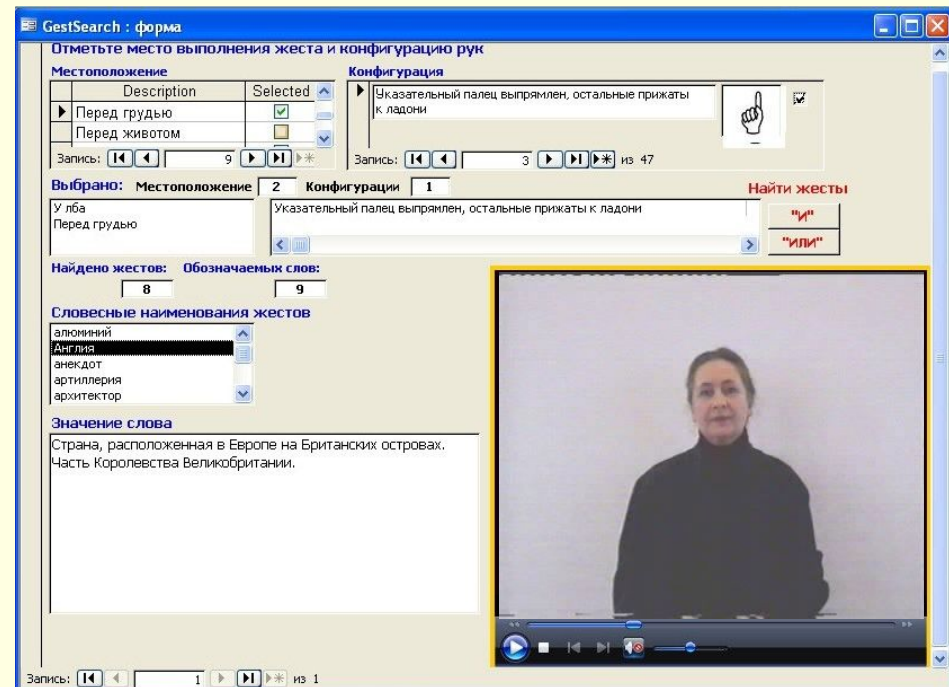
$$f_1: L \rightarrow G; f_2: K \rightarrow G$$

$$G_f = G_1 \cap G_2,$$

$$G_1 = f_1(\{x: x_1 \# x_2 \# \dots \# x_l \dot{Y} L_w\});$$

$$G_2 = f_2(\{x: x_1 \# x_2 \# \dots \# x_k \dot{Y} K_w\}).$$

В этом случае в результатах поиска будут отображены все жесты, содержащие хотя бы одну позицию и конфигурацию, заданные пользователем.



Эти методы поиска за счет уменьшения точности (выборка может включать жесты, лишь частично напоминающие искомые) полезны для целей обучения. За счет большей полноты поиска могут быть наглядно представлены результаты неточностей в исполнении жестов.

Разбиение жестового выражения на отдельные жесты

Многие жесты являются составными.

<i>Наименование жеста</i>	<i>Составляющие жесты</i>
<i>абитуриент</i>	<i>учиться, войти</i>
<i>аудитория</i> (помещение)	<i>учиться, комната</i>
<i>девочка</i>	<i>женщина,</i> <i>маленький</i>
<i>она</i>	<i>женщина, вы</i>

Разделение жестовой фразы на отдельные жесты следует вести, выбирая из словаря соответствующие по последовательности фаз жесты, имеющие наибольшую длину (число фаз), и анализируя семантику получаемого высказывания.

На значение жестового выражения влияет также многозначность жестов: одним и тем же жестом могут передаваться различные понятия и части речи, в частности существительные и глаголы.

При переводе текста в жесты следует учитывать, что некоторые понятия, передаваемые одним словом, передаются различными жестами в зависимости от пространственного положения объекта.

Понимание текста

- Результатом понимания текста должно быть опознание объектов, описанных в тексте, их пространственного положения, а также изменений их характеристик, действий и положения в соответствии с изменением времени текста.
- Опознание объектов включает в себя не только выделение именных групп, описывающих тот или иной объект, но и распознавание того, встречался ли ранее в тексте данный объект, совпадают ли объекты, имеющие одинаковые имена.

Система автоматической обработки текста должна уметь видеть?

- Разрешение дейктической референции при обработке жестовых высказываний требует знания о пространственном положении объекта.
- Анализ пространственного положения объектов, описываемых в тексте, подразумевает способность интеллектуальной системы представлять визуальные образы, сопоставлять их друг с другом, расставлять на пространственной шкале «ближе» – «дальше».

Спасибо за внимание

- Ответы на Ваши вопросы Вы можете получить на нашем стенде.