

Krassimir Markov, Vitalii Velychko, Oleksy Voloshin
(editors)

Natural and Artificial Intelligence

ITHEA

SOFIA

2010

Krassimir Markov, Vitalii Velychko, Oleksy Voloshin (ed.)

Natural and Artificial Intelligence

ITHEA®

Sofia, Bulgaria, 2010

ISBN 978-954-16-0043-9

First edition

Recommended for publication by The Scientific Council of the Institute of Information Theories and Applications FOI ITHEA

This book is engraved in prof. Zinovy Lvovich Rabinovich memory. He was a great Ukrainian scientist, co-founder of ITHEA International Scientific Society (ITHEA ISS). To do homage to the remarkable world-known scientific leader and teacher this book is published in Russian language and is concerned to some of the main areas of interest of Prof. Rabinovich.

The book is opened by the last paper of Prof. Rabinovich specially written for ITHEA ISS. Further the book maintains articles on actual problems of natural and artificial intelligence, information interaction and corresponded intelligent technologies, expert systems, robotics, classification, business intelligence; etc. In more details, the papers are concerned in: conceptual problems of the natural and artificial intelligent systems: structures and functions of the human memory, ontological models of knowledge representation, knowledge extraction from the natural language texts; network technologies; evolution and perspectives of development of the mechatronics and robotics; visual communication by gestures and movements, psychology of vision and information technologies of computer vision, image processing; object classification using qualitative characteristics; methods for comparing of alternatives and their ranging in the procedures of expert knowledge processing; ecology of programming – a new trend in the software engineering; decision support systems for economics and banking; systems for automated support of disaster risk management; and etc.

It is represented that book articles will be interesting for experts in the field of information technologies as well as for practical users.

General Sponsor: Consortium FOI Bulgaria (www.foibg.com).

Printed in Bulgaria

Copyright © 2010 All rights reserved

© 2010 ITHEA® – Publisher; Sofia, 1000, P.O.B. 775, Bulgaria. www.ithea.org; e-mail: info@foibg.com

© 2010 Krassimir Markov, Vitalii Velychko, Oleksy Voloshin – Editors

© 2010 Ina Markova – Technical editor

© 2010 For all authors in the book.

© ITHEA is a registered trade mark of FOI-COMMERCE Co.

ISBN 978-954-16-0043-9

C/o Jusautor, Sofia, 2010

МОДЕЛИРОВАНИЕ РЕАЛИСТИЧЕСКИХ ДВИЖЕНИЙ И МИМИКИ ДЛЯ ЗАДАЧ ВИЗУАЛИЗАЦИИ ЖЕСТОВОЙ ИНФОРМАЦИИ

Юрий Крак, Юрий Кривonos, Александр Бармак,
Антон Тернов, Богдан Троценко

Abstract: *Complex information technology for visualization communications gesture and mimicry is created.*

Keywords: *modelling, sign language, computer system.*

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

Введение

Задача исследования закономерностей и анимации движений является актуальной. Решение этой проблемы в виде создания инновационных технологий может существенно влиять на такие сферы человеческой деятельности как медицина, робототехника, компьютерная графика и системы виртуальной реальности.

Развитие современной науки, компьютеризация общества, использование мультимедийных и Интернет технологий создали достаточные условия для разработки компьютерных систем коммуникации для людей с дефектами слуха в формах и образах близких и понятных для них и для окружающего социума.

Люди с дефектами слуха для общения между собой используют жестовый язык. Основными визуальными средствами для воспроизведения этого языка есть фотографические и видео изображения жестов. Создавать с помощью этих средств современные учебные и коммуникационные компьютерные системы очень проблематично. Фотографическое изображение не дает нужной динамики жестов, а видео изображение достаточно громоздкое и в нем отсутствует нужная интерактивность (нельзя увидеть воспроизведение жеста с разных ракурсов, кроме того, в котором произведена видеозапись).

Эти существенные ограничения существующих средств воспроизведения жестового языка побуждают к разработке более гибких технологий, с помощью которых можно было бы создавать новые компьютерные системы обучения и коммуникации для людей с дефектами слуха. В развитие этого предложена концепция [Кривonos, 2008, 2009] информационной технологии невербального общения людей с дефектами слуха. Комплексная информационная технология включает в себя кроме другой и функциональность по синтезу движений жестового языка и дактильной азбуки на трехмерной модели человека.

Реализация возможности генерации анимации процесса общения с помощью жестового языка с использованием виртуальных моделей людей требует разработки соответствующих информационных и математических моделей. Исходя из этого, сформулирована следующая постановка задачи:

- необходимо разработать информационную и математическую модель для фиксации морфем (минимально значимых единиц) жестового языка;
- в рамках этой модели необходимо разработать технологию и соответствующее программное обеспечение для получения, сохранения и воспроизведения жестов;
- необходимо предложить алгоритмическое решение для расчета человекоподобных траекторий движения рук и корпуса модели при переходах от жеста к жесту.

Модель для фиксации единиц жестового языка

Процесс воспроизведения жеста трехмерной моделью человека можно считать анимацией с соответствующей частотой разных состояний упрощенной скелетной модели человека.

Скелетная модель человека упрощенно воссоздает скелет живого человека. Ее можно формализовать как иерархическую структуру, состоящую из соединенных кинематических пар, которые воссоздают основные кости человеческого скелета. Современные пакеты трехмерного моделирования (Poser, 3D Studio Max) умеют генерировать анимацию с помощью виртуальной статической модели и информации о изменении соответствующих углов скелета. Для формального описания процесса фиксации жеста можно использовать множество, отображающее упрощенный скелет человека (рис. 1) и изменения значений углов Эйлера и порядка их применения для соответствующих костей этого скелета с течением времени (дискретно, с соответствующей частотой (1/30 сек, например):

$$H = \{H_i : H_i = \{k, d_i, M_i \in M\}\}, \quad (1)$$

где H_i – i -я кость скелета ($i = 0, \dots, N - 1$, N – количество костей в скелете); k – индекс кости-предка; $d_i = [x_i, y_i, z_i]^T$ – координаты точки – конца кости в системе координат, связанной с началом этой кости;

$$M = \{M_i : M_i = (order_i, \theta_i)\}, \quad (2)$$

где M_i – для i -й кости значения углов Эйлера и порядок применения поворотов для кости с течением времени; $order_i \in \{1, \dots, 6\}$ – порядок применения поворотов вокруг соответствующих координатных осей для i -й кости (1-XYZ, 2-XZY, 3-YXZ, 4-YZX, 5-ZXY, 6-ZYX); $\theta = (\theta^j)$, $\theta^j = \{\theta_i^j : \theta_i^j = (\varphi_{iX}^j, \varphi_{iY}^j, \varphi_{iZ}^j)\}$ – множество изменений углов Эйлера для i -й кости с течением времени ($j = 0, \dots, K - 1$, K – количество кадров для воспроизведения движения с заданной частотой).

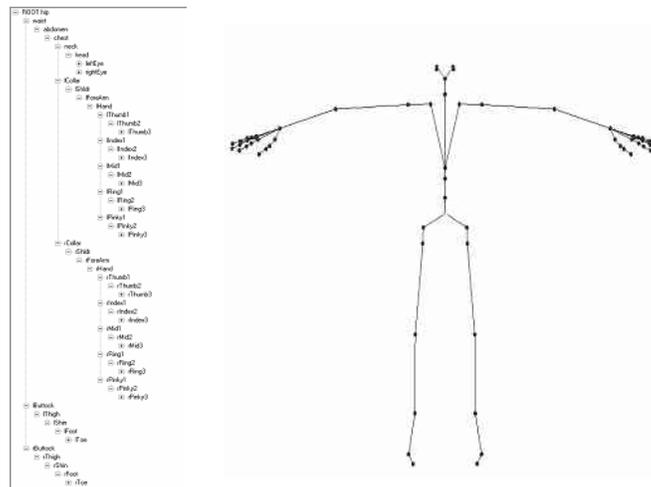


Рис.1. Упрощенный скелет человека

Технология получения и сохранения жестов

Для получения множества углов, характеризующих изменения положения костей от начального состояния скелета предложено использовать технологию захвата движения (motion capture).



Рис.2. Приложения для видеозахвата движения.

Создано соответствующее программное обеспечение (рис. 2), с помощью которого возможно оцифровывать жесты.

Синтез процесса анимации жестового языка

Для синтеза процесса анимации жестов и мимики трехмерной моделью человека предложено следующее формальное описание, использующее соответствующие множества параметров и алгоритмы работы с ними. Трехмерная модель человека, на которой будет реализоваться процесс анимации жестов и мимики, имеет следующие атрибуты: $V = \{v_i : v_i = (x, y, z)\}$ - множество вершин треугольников для триангуляции поверхности трехмерной модели человека; $N = \{n_i : n_i = (x, y, z)\}$ - множество нормалей к вершинам; $T = \{t_i : t_i = (t, s)\}$ - множество текстурных координат к вершинам; $V^{ind} = \{V_i^{ind} : V_i^{ind} = (k_1, k_2, k_3)\}$ - множество индексов, указывающих порядок построения треугольников из множества вершин; $I = \{I_i : I_i = \{img\}\}$ - множество фотографических изображений элементов модели – текстуры.

Для моделирования скелетной анимации необходимо уметь рассчитывать новые значения вершин треугольников (V). Для этого предлагается использовать механизм скининга. Скининг можно определить как алгоритм привязки множества вершин треугольников, определяющих поверхность модели к значениям углов скелета. Тогда модель скелетной анимации можно формализовать следующим образом:

$MH = \{MH_i : HM_i = \{k, \{l_1, \dots, l_m\}, d_i, Glb_i, Order_i\}\}$ - описание упрощенного скелета человека (иерархия костей) для реализации скелетной анимации, где MH_i - i -я кость скелета ($i = 0, \dots, N - 1$, N - количество костей в скелете); k - индекс кости-предка; $\{l_1 \dots l_m\}$ - множество индексов дочерних костей, $d_i = [x_i, y_i, z_i]^T$ - координаты точки - конца кости в системе координат, связанной с началом этой кости; Glb - вектор для определения координат кости в глобальной системе координат, $Order_i$ - порядок применения поворотов.

$Skin = \{Skin_i : Skin_i = \{(IndexVertex_1, Weight_1), \dots\}\}$ - множество вершин, влияющих на текущую вершину при изменении углов.

Скининг рассчитывается для каждой вершины V следующим образом:

$$v'_j = \sum_{i=0}^N \left\{ (v_j * IBM_{H_i} * JM_{H_i}) * JW_{H_i} \right\}, \quad (3)$$

где: n - количество костей, связанных с вершиной v ; IBM_{H_i} - инверсная bind-pose матрица для кости H_i ; JM_{H_i} - матрица перемещения для кости H_i ; JW_{H_i} - весовой коэффициент для применения влияния точек кости H_i на вершину v .

Для моделирования анимации проговаривания и эмоционального окрашивания предложено использовать механизм морфинга. Морфинг можно определить как алгоритм плавного перехода от одного состояния объекта к другому. При морфинге используются только опорные состояния, с помощью которых рассчитываются промежуточные состояния и моделируется процесс анимации. Модель морфинга мимики проговаривания можно формализовать следующим образом:

Воспроизведение или построение мимики на лице трехмерной модели достигается при применении относительного (сегментного) морфинга к модели головы и мимического процесса. Формула относительного морфинга для M морфов в формализме модели имеет следующий вид:

$$V' = V + \sum_{m=1}^M \omega_m \cdot TM_m, \quad (4)$$

где ω_m - весовые коэффициенты, V - меш базовой модели без морфинга, TM_m - входной меш (морф) для блендинга (blending). Результатом операции есть линейная комбинация мешей модели и мимики.

Метод сегментного морфинга дает возможность: формировать несколько выражений лица на основе небольшого количества морфов и изменять состояние лица при анимации проговаривания. Дополнительным преимуществом сегментного морфинга есть то, что можно анимировать челюсть независимо от губ и глаз (моргание), независимо от эмоциональных проявлений на лице.



Рис. 3. Морфинг-мишени визем и эмоций

В работе предлагается осуществить синтез посредством морфинга визем украинской речи (рис. 3) в процессе проговаривания губами передаваемой словесной информации. Морфинг с использованием морфинг-мишеней визем осуществляется с учетом весовых коэффициентов:

- определяется количество кадров для анимации;
- определяется набор эмоций присутствующих в слове;
- по фонетической транскрипции рассчитывается набор визем для визуализации процесса артикуляции;

- для визем v_i рассчитываются длительности этапов анимации и точки появления виземы – минимальный номер кадра в котором $w_{v_i} > 0$.
- Для синхронизации анимации жеста и мимики полагалось, что начало и конец анимации мимики артикуляции должны совпадать с началом и концом анимации жеста.

Реализация воспроизведения процесса анимации

Для воспроизведения процесса анимации жестов и мимики трехмерной моделью человека создано приложение, реализующее скелетную (для воспроизведения жеста) и морфемную (для воспроизведения мимики проговаривания и эмоций) анимацию. Реализованная соответствующая программная функциональность, которая, используя трехмерное API OpenGL, воспроизводит по заданным атрибутам модель человека (рис. 4). Модель, с использованием алгоритмов скининга и морфинга, воспроизводит анимацию (рис. 5).



Рис. 4. Трехмерная модель человека



Рис. 5. Кадры анимации жеста «Ты»

Для реализации украинского жестового языка создано приложение (рис. 6), которое воссоздает методику преподавания жестового языка в специализированных общеобразовательных школах для глухих детей.

Функционально приложение состоит из трех информационных блоков (темы, слова и предложения) и блока воспроизведения жеста виртуальной моделью.

Блок воспроизведения жеста виртуальной моделью несет особенную функцию. С помощью него стало возможным демонстрировать в учебном процессе динамику жеста. Беря во внимание то, что жесты оцифровываются носителями жестового языка, они, по сути, делаются стандартами воспроизведения жеста. Реализованная в блоке функциональность для покадрового показа жеста служит средством, с помощью которого станет возможным изучение жеста без особенностей показа конкретным преподавателем. Это станет основой того, что изученные детьми жесты будут одинаковыми для разных школ, регионов. Фактически, жестовый язык (в основном своем множестве) станет стандартным на территории Украины.



Рис. 6. Приложение «Украинский жестовый язык».

Выводы

Используя модель для фиксации движений, которые воспроизводят украинский жестовый язык, было оцифровано множество из 50 жестов. Воспроизведение жестов из этого множества (с помощью трехмерной модели) показало способность предложенной технологии весьма реалистично воспроизводить на трехмерной модели движения, полученные с видео-изображения конкретного человека – носителя жестового языка.

Реализовано приложение, воспроизводящее уроки украинского жестового языка.

Созданная технология и программное обеспечение может стать основой создания стандарта жестового языка. Станет возможным решить проблему отличий для одних и тех самых жестов, которая возникает от того, что дети изучают жест, который содержит особенности конкретного преподавателя.

Дальнейшие исследования направлены на усовершенствование предложенной технологии:

- наполнения базы данных жестов основным множеством жестов украинского жестового языка – создание стандарта жестового языка;
- создание средства для семантического связывания предложений украинского языка с предложениями на жестовом языке.

Благодарности

Статья частично финансирована из проекта ITHEA XXI Института Информационных теорий и Приложений FOI ITHEA и Консорциума FOIBulgaria (www.itea.org, www.foibg.com).

Библиография

- [Кривонос, 2008] Кривонос Ю.Г., Крак Ю.В., Бармак О.В., Тернов А.С. Информационная технология невербального общения людей с проблемами слуха // Искусственный интеллект. – 2008. №3. – ст. 325-331
- [Кривонос, 2009] Кривонос Ю.Г., Крак Ю.В., Бармак О.В., Тернов А.С., Троценко Б.А. Информационная технология для моделирования украинского жестового языка // Искусственный интеллект. – 2009. №3. – ст. 186-198

Authors' Information



Iurii Krak – *V.M.GlushkovCybernetics Institute of NASU, senior scientist*
krak@unicyb.kiev.ua



Iurii Kryvonos - *V.M.GlushkovCybernetics Institute of NASU, Director Deputy,*



Alexander Barmak – *V.M.GlushkovCybernetics Institute of NASU, senior scientist,*
barmak@svitonline.com



Anton Ternov – *V.M.GlushkovCybernetics Institute of NASU, junior scientist,*
anton.ternov@gmail.com



Bohdan Trotsenko– *V.M.GlushkovCybernetics Institute of NASU, junior scientist,*
bohdan@trotsenko.com.ua

address: 40 Glushkov ave., Kiev, Ukraine, 03680