



Ключевые слова: *жестовая речь, жестовая нотация, моделирование, распознавание жестовой речи.*

ВВЕДЕНИЕ И ПОСТАНОВКА ЗАДАЧИ

Рассматриваемая в данной статье жестовая речь (ЖР) является средством общения людей с недостатками слуха как между собой, так и с другими людьми, в том числе слышащими. В ЖР для передачи информации используются визуально-кинестические средства (жесты рук, артикуляция губ, мимические проявления эмоций). Например, украинская ЖР сочетает в себе несколько разновидностей жестовой коммуникации: дактильную азбуку для передачи букв украинского алфавита, а также для передачи слов, языковых конструкций и семантических концептов. Дактильная азбука (воспроизводящая буквы с помощью пальцев) выполняет вспомогательную функцию (для передачи имен, названий и т.д.). В ЖР каждому смысловому понятию соответствует определенный уникальный жестовый эквивалент. Система ЖР подразделяется на две подсистемы: непосредственно ЖР и калькирующую ЖР. Калькирующая ЖР полностью копирует разговорный язык.

Основное, что объединяет все жестовые системы общения, — это способ передачи и восприятие информации. Базовой смысловой единицей здесь выступает жест (визуально-кинестический акт), в котором задействованы руки, мимика лица и артикуляция губ. Кинестическая природа жеста и его визуальное восприятие предопределяют особенности ЖР: возможность определенным образом расположить жест в пространстве, выполнить одновременно два жеста двумя руками и т.д.

Специфика украинской ЖР (как и русской, белорусской) заключается в том, что диалекты разных регионов несколько отличаются между собой. Современные средства вычислительной техники, WEB-технологий позволят преодолеть эти диалектические различия. Для этого необходимы современные технологии — мультимедийные компьютерные средства для описания и воспроизводства ЖР. Для украинской ЖР таких средств нет. Поэтому актуальной проблемой является создание современных технологий нотации и воспроизведения ЖР.

Лексические жестовые элементы и в самой ЖР, и в дактильной азбуке формируются практически одинаково и могут быть форматизированы следующим образом: по конфигурации (фигура руки или рук); по месту исполнения (место положения руки в момент выполнения жеста); по характеру движения (какие движения выполняют руки). Такое описание впервые предложено в работе [1] для американской ЖР (AMSLEN).

Для практического применения (включая компьютерное) ученые разработали несколько систем нотации, являющихся наглядными и в то же время унифицированными для быстрого ввода жестов в словарь. Одна из наиболее известных систем нотации — SignWriting [2], знаки в которой символизируют мимику, движения рук, тела и пишутся сверху вниз. Эта система отличается большей проработанностью и однозначностью записи, однако у нее чрезвычайно большой набор знаков (более тысячи), что делает ее малоприменимой для применения в компьютерных технологиях (даже с использованием возможностей Unicode-кодирования).

На практике более предпочтительно использовать меньшее количество отдельных случаев конфигурации руки, описывая их как переменные (например, количество задействованных пальцев) при одной константе (форме ладони). Такому требованию, например, удовлетворяет система нотации жестов, предложенная в [3]. В этой нотации выделяется более 30 конфигураций рук, около 50 характеристик места выполнения жеста и более 70 характеристик локализации.

За последние годы широкое распространение получила гамбургская система нотации HamNoSys (Hamburg Notation System) [4], разработанная в конце XX века в Германии. Она отличается наибольшей проработанностью и пригодна для использования в компьютерных программах, так как ее знаки переведены в кодировку Unicode с соответствующими компьютерными шрифтами. Инвентарь HamNoSys позволяет записать практически любой жест, что делает эту систему универсальной и пригодной для любой ЖР мира.

Одним из наиболее эффективных средств обучения и человеко-машинного взаимодействия являются мультимедийные компьютерные программы, поэтому создание информационных программ, способных работать с ЖР (синтезировать и распознавать), является одной из приоритетных задач при работе с неслышащими и слабослышащими людьми. Возможным вариантом компьютерного синтеза ЖР является использование анимированного аватара человека, который управляется с помощью символов жестовой нотации, описывая необходимые конфигурации рук и различные типы движений. Жесты из лексикона в такой системе синтеза представляют цепочку символов в выбранной нотации, поэтому словарь жестов может легко модифицироваться и расширяться без использования специального оборудования.

Для украинской ЖР предложена концепция [5] информационной технологии невербального общения людей с недостатками слуха. Концепция заключается в использовании мультимедийных возможностей современной компьютерной техники и систем коммуникаций, а именно использовании пространственных моделей людей для воспроизведения ЖР. В рамках этой концепции предложены математическая модель и ее компьютерная реализация для получения и воспроизведения ЖР с помощью трехмерных моделей людей. Для реализации возможностей генерации анимации процесса вещания с помощью ЖР с использованием виртуальных моделей людей были разработаны следующие информационные и математические модели [6]:

- 1) информационная и математическая модели для фиксации морфем (минимально значимых единиц) жестового языка;
- 2) технология и соответствующее программное обеспечение для получения, сохранения и воспроизведения жестов.

Предложенная технология моделирования движений виртуального персонажа для пространственного воспроизведения жестового языка помимо явных преимуществ перед другими системами содержит существенный недостаток: необходимость оцифровки всего множества жестов с помощью носителя ЖР для каждой новой модели. Ресурсоемкость такого подхода как по времени, так и по

стоимости побуждает к поиску других подходов, учитывающих такие ограничения. Кроме того, существование систем нотаций для ЖР [2–4] побуждает к созданию такой модели формализации ЖР, которая позволила бы описывать жесты в этих системах. Цель настоящей статьи — разработка такой модели формализации для украинской ЖР.

Сформулируем следующую постановку задачи: предложить модель для формализации жестовых единиц с целью:

- описания жестовых единиц в нотации предложенной модели;
- создания произвольного жеста на основе фиксированных множеств ключевых состояний, траекторий и т.д. с возможностями его распознавания.

МЕТОД ПРОЕКТИРОВАНИЯ ДВИЖЕНИЙ ВИРТУАЛЬНОГО ПЕРСОНАЖА

Процесс воспроизведения жеста 3D моделью человека можно считать анимацией с соответствующей частотой различных состояний скелетной модели человека. Скелетная модель человека упрощенно воспроизводит скелет живого человека. Ее можно формализовать как иерархическую структуру, состоящую из соединенных кинематических пар, которые воспроизводят структуру основных костей скелета человека. Современные пакеты трехмерного моделирования (Poser, 3D Studio Max) позволяют генерировать анимацию с помощью виртуальной статической модели и информации об изменении соответствующих углов сочленений скелета.

Для проектирования движений анимированного персонажа в задачах пространственного воспроизведения наиболее часто применяется схема представления виртуального персонажа в виде дерева. Суть этой схемы заключается в моделировании скелета персонажа с помощью дерева $S = D(V, E)$, где V — множество звеньев скелета, E — множество соединений (ребер дерева). Такое дерево называют кинематической схемой [7]. За основу принимают узел, совпадающий с центром масс персонажа, с которым соединены ветви рук, ног и т.д. С каждым звеном связываются локальная система координат и матрица перехода от предыдущего звена к текущему звену. Матрицы перехода зависят от значений обобщенных координат, определяющих связи между звеньями, входящими в данную кинематическую пару. Для основного (корневого) узла скелета матрица перехода определяет переход из глобальной системы отсчета в систему отсчета персонажа. Такой подход удобен для описания движения персонажа и используется в системах и методах анимации. В наиболее известных методах моделирования движений, используемых в современных системах 3D графики, процесс движения персонажа может быть представлен с помощью семейства функций $m1(t), m2(t), \dots, mn(t)$ — изменений значений обобщенных координат за время t , где n — число обобщенных координат кинематической цепи скелета. Кроме этого, процесс движения персонажа может быть описан на основе методов моделирования физики и биомеханики [8, 9]. Тогда движения персонажа могут быть представлены семейством функций $f1(t), f2(t), \dots, fn(t)$ — изменением значений напряжений в мышцах. Но такой подход требует большого количества параметров, описывающих динамическую схему персонажа. К преимуществам такого подхода можно отнести уменьшенный объем информации о движении и использовании физических свойств персонажа в процессе движения.

Для полноценного воспроизведения жестового языка предлагается использовать модель упрощенного скелета человека, используя 56 звеньев (костей). Особая детализация сделана в области рук ввиду их основной нагрузки по реалистичному воссозданию жестов — задействованы почти все кости скелета человека. С каждым звеном связана своя евклидова система координат. Поворот од-

ного звена относительно другого характеризуется параметрами изменения положения данного звена относительно предыдущего при условии, что система координат предыдущего звена фиксирована. Эти параметры являются обобщенными координатами. Положение всего скелета исчисляется в фиксированной (абсолютной) системе координат.

В качестве обобщенных координат предложено использовать углы Эйлера. Такой выбор обусловлен тем, что подавляющее большинство систем захвата движений продуцирует информацию о движении в виде BVH файлов, которые используют углы Эйлера [10]. Следовательно, для формального описания процесса фиксации жеста можно использовать множество, содержащее упрощенный скелет человека, значения углов Эйлера и порядок их применения для соответствующих костей этого скелета с течением времени (дискретно, с соответствующей частотой (1/30 сек) и т.д.).

Для получения множества углов, воспроизводящих процесс анимации жестов, использовалась технология захвата движения (motion capture) [10,11]. Множество углов изменения движения считывалось с человека-носителя ЖР. Это позволило получить достаточно реалистичные движения.

Множества углов ($\theta = (\theta^j)$, $\theta^j = \{\theta_i^j: \theta_i^j = (\varphi_{iX}, \varphi_{iY}, \varphi_{iZ})\}$), характеризующих изменения положения костей относительно исходного состояния скелета, определялись по следующему алгоритму.

1. С помощью трех камер проводилась съемка человека, воспроизводящего конкретный жест. Камеры устанавливались на одинаковом фокусном расстоянии (L) от объекта съемки и фиксировали проекции — фронтальную, левую и правую.

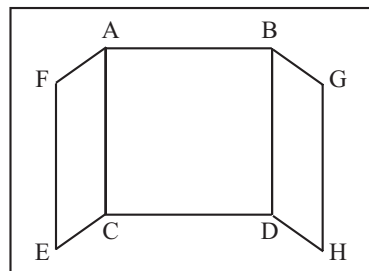


Рис. 1

2. Для определения углов в сочленениях скелета для конкретного состояния воспроизводились соответствующие кадры съемки в трех плоскостях, расположенных на заднем плане трехмерной сцены (рис. 1). Скелет располагался на фоне поверхностей. На поверхности проецировались кости под определенными углами, которые необходимо определить.

Полагаем, что объем сцены ограничен кубом с координатами точек вершин по диагоналям $(-1; -1; -1)$ и $(1; 1; 1)$. Тогда плоскости для отображения кадров съемки проекций костей будут ограничены координатами (здесь обозначено $c() = \cos()$): $A(-0,5; 0,5; -1)$, $B(0,5; 0,5; -1)$, $C(-0,5; -0,5; -1)$, $D(0,5; -0,5; -1)$, $F(-c(45^\circ); 0,5; -1 + c(45^\circ))$, $E(-c(45^\circ); -0,5; -1 + c(45^\circ))$, $G(c(45^\circ); 0,5; -1 + c(45^\circ))$, $H(c(45^\circ); -0,5; -1 + c(45^\circ))$. Координаты точек для построения проекций костей на эти плоскости получаем следующим образом:

- точки скелета поворачиваем вокруг оси Y на 180° (чтобы скелет «смотрел» на фронтальную плоскость): $(x_i^{180^\circ}, y_i^{180^\circ}, z_i^{180^\circ}) = (x_i, y_i, z_i)R_y$, где R_y — матрица поворота на 180° вокруг оси Y ;
- для плоскости, ограниченной квадратом с вершинами A, B, D, C , точки сочленения костей скелета $(x_i^{180^\circ}, y_i^{180^\circ}, z_i^{180^\circ})$ проецируются в точки $(x_i^{180^\circ}, y_i^{180^\circ}, -1)$;
- для плоскостей, ограниченных прямоугольниками $FACE$ и $GBDH$, исходя из того, что они расположены под углом 45° к фронтальной плоскости, точки

сочленения костей скелета $(x_i^{180^\circ}, y_i^{180^\circ}, z_i^{180^\circ})$ проецируются в точки $(x_i^*, y_i^{180^\circ}, z_i^*)$. Учитывая, что прямая, проходящая через точку $(x_i^{180^\circ}, y_i^{180^\circ}, z_i^{180^\circ})$ и перпендикулярная прямой $z = ax + b$, описывается уравнением $z = -(x - x_i) / a + z_i$, точка пересечения этих прямых (т.е. точка проекции) определяется как $x_i^* = (az_i + x_i - ab) / (a^2 + 1)$, $z_i^* = (a^2 z_i + ax_i + b) / (a^2 + 1)$. Здесь для левой проекции уравнение прямой, проходящей через точки $(x_1, z_1) = (-c(45^\circ), -1 + c(45^\circ))$ и $(x_2, z_2) = (-0.5, -1)$, имеет вид $(z - z_1) = (z_2 - z_1)(x - x_1) / (x_2 - x_1)$. Отсюда $z = -3,414211x - 2,707106$, т.е. $a = -3,414211$ и $b = -2,70711$. Для правой проекции точек $(c(45^\circ), -1 + c(45^\circ))$ и $(0,5; -1)$ имеем $a = 3,414211$ и $b = -2,70711$.

3. Сопоставляем точки проекций костей скелета с соответствующим им изображением реального человека (для фронтального, левого и правого кадров), т.е. получаем новые координаты: $(x_i^{\text{front}}, y_i^{\text{front}}, -1)$ — для фронтальной проекции, $(x_i^{\text{left}}, y_i^{180^\circ}, z_i^{\text{left}})$ — для левой проекции и $(x_i^{\text{right}}, y_i^{180^\circ}, z_i^{\text{right}})$ — для правой проекции, где

$$\begin{aligned} x_i^{\text{front}} &= x_i^{180^\circ} + \text{off}_i^{\text{front}_x}, \quad y_i^{\text{front}} = y_i^{180^\circ} + \text{off}_i^{\text{front}_y}, \quad x_i^{\text{left}} = x_i^* + \text{off}_i^{\text{left}_x}, \\ z_i^{\text{left}} &= z_i^* + \text{off}_i^{\text{left}_z}, \quad x_i^{\text{right}} = x_i^* + \text{off}_i^{\text{right}_x}, \quad z_i^{\text{right}} = z_i^* + \text{off}_i^{\text{right}_z}. \end{aligned} \quad (1)$$

Полученные таким образом значения смещений *off* используются для всех кадров, отображающих жест.

4. Для каждого кадра на паре изображений (фронтальное и левое или фронтальное и правое) указываем точки сочленений костей, которые изменили положение $((x_{i,\text{new}}^{\text{front}}, y_{i,\text{new}}^{\text{front}})$ и $(x_{i,\text{new}}^{\text{left or right}}, y_{i,\text{new}}^{\text{left or right}}))$, и определяем их трехмерные координаты (в системе координат, связанной с корневой точкой скелета):

$$\begin{aligned} (x_i^{\text{new}}, y_i^{\text{new}}, z_i^{\text{new}}) &= \\ &= (x_{i,\text{new}}^{\text{front}} - \text{off}_{i,x}^{\text{front}}, y_{i,\text{new}}^{\text{front}} - \text{off}_{i,y}^{\text{front}}, x_{i,\text{new}}^{\text{left or right}} - \text{off}_{i,x}^{\text{left or right}}). \end{aligned} \quad (2)$$

Имея трехмерные координаты нового положения точек сочленения костей, определим значения углов Эйлера, на которые поворачивают каждую кость из предыдущего состояния в текущее. Для этого необходимо решить следующую задачу. Пусть в системе координат, в которой имеются два единичных вектора: $r_1 = (x_1, y_1, z_1)$ и $r_2 = (x_2, y_2, z_2)$, необходимо найти углы Эйлера $\varphi_x, \varphi_y, \varphi_z$ вращения вокруг соответствующих осей OX, OY, OZ , которые переводят вектор r_1 в вектор r_2 . Следует учесть, что умножение матриц не коммутативно, поэтому будем рассматривать шесть возможных вариантов, характеризующих порядок вращения вокруг осей: $OZYX, OYXZ, OYZX, OXYZ, OXZY, OZXY$.

Углы Эйлера, например, когда порядок вращения ZYX определяются по формулам (здесь обозначено $a() = \text{arctg}()$), имеют вид

$$\begin{aligned} \varphi_x &= a(t_{32} / t_{33}) + n\pi, \quad \varphi_y = a\left(-t_{31} / \sqrt{t_{11}^2 + t_{21}^2}\right) + n\pi, \\ \varphi_z &= a(t_{21} / t_{11}) + n\pi, \end{aligned} \quad (3)$$

где

$$t_{11} = x_1x_2 + y_1y_2 + z_1z_2, \quad t_{21} = -y_1x_2 + x_1y_2 / d_1 + y_1z_1z_2 / d_1, \quad t_{31} = z_1x_2 + z_2d_1, \\ t_{32} = -z_1y_2 + y_2z_2d_1 / d_2, \quad t_{33} = z_1z_2 + d_1d_2, \quad d_1 = \sqrt{x_1^2 + y_1^2}, \quad d_2 = \sqrt{x_2^2 + y_2^2}.$$

Для других порядков вращения углы определяются аналогично.

Для синтеза процесса анимации жестов и мимики трехмерной моделью человека было предложено следующее формальное описание, использующее соответствующие множества параметров и алгоритмы работы с ними. Трехмерная модель человека, на которой реализуется 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; $V^{\text{ind}} = \{V_i^{\text{ind}}: V_i^{\text{ind}} = \{k_1, k_2, k_3\}\}$ — множество индексов, которые указывают порядок построения треугольников из множества вершин; $I = \{I_i: I_i = \{img\}\}$ — множество фотографических изображений элементов модели — текстуры.

Для формализации предлагается разложить жест на следующие составляющие, которые его однозначно характеризуют: 1) конфигурация пальцев руки, 2) локализация выполнения, 3) траектория (направление) и характер движения.

ПОСТРОЕНИЕ ДВИЖЕНИЙ И РАСПОЗНАВАНИЕ ДАКТИЛЬНОЙ ИНФОРМАЦИИ





Рассмотрим более детально основные конфигурации пальцев руки человека, которые зависят от ключевых положений каждого пальца и ладони. Итоговая конфигурация будет суперпозицией конкретных значений положений. Представим конфигурацию пальцев следующим множеством:

$$F = \left\{ \begin{array}{l} \{F_i^k(t): F_i^k(t) = \{L^{\text{beg}}, L^{\text{end}}, t\}\}, \\ \{F_i^k: F_i^k = \{L^{\text{beg}}\}\}, \end{array} \right. \quad (4)$$

где F_i^k — i -я конфигурация k -го пальца ($i = \overline{0, p}$, $k = \overline{1, 5}$); L^{beg} , L^{end} — множества названий костей и значений углов Эйлера в этих костях соответственно для начального и конечного состояний конфигурации пальца; t — интерполяционный коэффициент для расчета промежуточного состояния ($t \in [0, 1]$), p — некоторое фиксированное положительное число.

В табл. 1 даны описания показателей возможных конфигураций большого пальца руки ($k = 1$). Для остальных пальцев определены следующие конфигурации: для указательного пальца ($k = 2$): $F_0^2(t)$ — от прижатого состояния к выпрямленному, F_1^2 — вытянут в горизонтальном положении, F_2^2 — указательный и большой разведены; для среднего пальца ($k = 3$): $F_0^3(t)$ — от прижатого состояния к выпрямленному, F_1^3 — вытянут в горизонтальном положении, F_2^3 — указательный и средний разведены, F_3^3 — средний лежит на прямом указательном; для безымянного пальца ($k = 4$): $F_0^4(t)$ — от прижатого состояния к выпрямленному, F_1^4 — вытянут в горизонтальном положении, F_2^4 — средний и безымянный разведены; для мизинца ($k = 5$): $F_0^5(t)$ — от прижатого состояния к выпрямленному, F_1^5 — вытянут в горизонтальном положении, F_2^5 — безымянный и мизинец разведены.

Таблица 1

$F_i^k(t)$	Описание конфигураций пальцев	Координаты пальцев и их положение			
$F_0^1(t)$	От прижатого состояния ($t = 0$) к выпрямленному ($t = 1$)	Thumb1 [3.5, -7.9, -5.7] Thumb2 [17, -19.5, 28.5] Thumb3 [24.1, -28.2, 0.5]		Thumb1 [10.1, 20.3, 7.9] Thumb2 [3.5, 6.4, -8.9] Thumb3 [3.7, 30.5, -8.7]	
F_1^1	Большой палец соединен подушечкой с указательным	Thumb1 [37.1, 7.9, -10.8] Thumb2 [22.8, 8.6, 4.6] Thumb3 [23.4, 0.5, -15.05] Index1 [26.4, -10.4, 39.3] Index2 [26.8, 7.8, 31.2] Index3 [26.8, 7.8, 31.2]			
F_2^1	Большой палец прижат к ладони	Thumb1 [36.7, -39.6, 19.4] Thumb2 [23.6, -23.5, 5.05] Thumb3 [19.03, 1.3, -1.7]			

Для указания конфигурации пальцев руки, кроме конфигураций самих пальцев, определим положение ладони. Положение ладони однозначно определяется соответствующими значениями углов Эйлера:

$$D = \{D_i: D_i = \{\varphi_i^x, \varphi_i^y, \varphi_i^z\}\}. \quad (5)$$

Определим следующие конкретные положения ладони: D_1 — вверх, D_2 — стандартное, D_3 — вниз, D_4 — к себе, D_5 — в сторону. На рис. 2 показаны некоторые положения ладони, необходимые для описания дактильной азбуки.

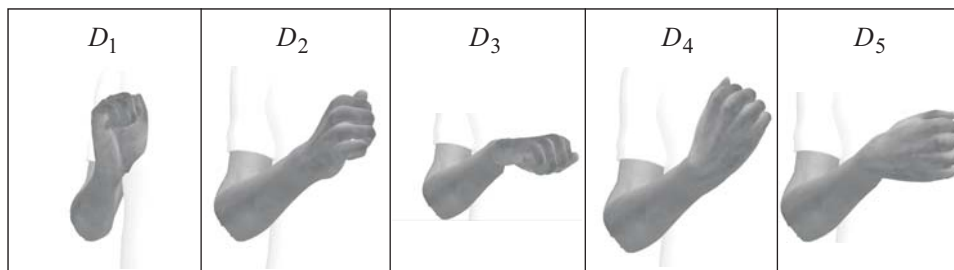


Рис. 2

Для локализации выполнения жеста фиксируются состояния костей скелета для областей, находящихся на пересечении уровней горизонтальных (над макушкой головы, на уровне лба, линия бровей, линия глаз, уровень правого и левого уха, уровень щек, рта, подбородка, под подбородком, линия шеи, плечи, грудь, талия, линия ниже талии) и вертикальных (слева, в центре, справа). Для каждой локализации указывается признак — с прикосновением или без прикосновения:

$$L = \{L_i(\mu): L_i(\mu) = \{K_i, \mu\}\}, \quad (6)$$

где L_i — состояние, отвечающее за конкретную локализацию; K_i — названия костей и значения углов Эйлера для них; μ — признак касания (1) или не касания (0). Опишем изложенным способом несколько знаков дактильной азбуки (табл. 2).

Таблица 2

Буква	Формула для определения положения руки и ее описание	Технология motion capture	По формуле
В	$R(L_1(0), D_1, F_0^1(0.5), F_0^2(1), F_0^3(1), F_0^4(1), F_0^5(1));$ правая рука: в районе груди, справа, без касания, ладонь в сторону все пальцы выпрямлены		
Г	$R(L_1(0), D_5, F_0^1(1), F_0^2(1), F_0^3(0), F_0^4(0), F_0^5(0));$ правая рука: в районе груди, справа, без касания, ладонь вверх, большой палец выпрямлен наполовину, остальные пальцы в прижатом состоянии		

Предложенная формализация описания конфигурации рук позволяет разрабатывать методы распознавания произвольных единиц жестового языка. Рассмотрим, например, распознавание дактилем. Поскольку имеется множество ключевых конфигураций (4), то в качестве вектора характеристических признаков эффективно использовать геометрические параметры этих конфигураций: контур кисти руки, ее центр масс, компактность, ориентацию, растяжение, моментные значения и т.п. [12]. Важным для задач распознавания является использование контурного анализа изображения, которое позволяет задачу распознавания свести к задаче идентификации элементов дактильного языка. Отметим, что при распознавании дактилем контурный анализ позволяет учитывать разные размеры руки на разных фокусных расстояниях. Для проверки эффективности использования геометрических признаков изображения для распознавания дактилем были проведены эксперименты, в которых принимали участие люди разного возраста и с разными размерами руки. В процессе эксперимента человек показывал определенную дактилему 50 раз на разных фокусных расстояниях, при этом подсчитывалось количество правильно распознанных дактилем. Результаты работы системы по распознаванию букв дактильной азбуки жестового языка продемонстрировали, что при использовании предложенных характеристических признаков можно осуществить распознавание на уровне 80%, что подтверждает эффективность предложенного подхода.

ПОСТРОЕНИЕ ТРАЕКТОРИЙ-ПЕРЕХОДОВ МЕЖДУ ЖЕСТОВЫМИ ЭЛЕМЕНТАМИ

В работах [5, 6] дана технология для синтеза промежуточных траекторий-переходов между жестовыми единицами в конструкциях-предложениях. В рамках используемой модели был найден функциональный переход от конечного состояния упрощенного скелета (начальный вектор вектор обобщенных координат — значения углов вращения вокруг соответствующих осей для каждой кости) одного жеста к исходному состоянию (другой вектор обобщенных координат) иного жеста. Найденный функциональный переход породил траекторию, визуально близкую к траектории движения человека при аналогичном переходе.

Для формализации жеста и с целью дальнейшего применения механизма агглютинации (для создания жестовых конструкций типа предложений) было использовано множество

$$G = \{G_i: G_i = \{f_i^{\text{beg}}, f_i^{\text{end}}, S_i^{\text{beg}}, S_i^{\text{end}}, \theta_i\}\}, \quad (7)$$

где G_i — конкретный жест ($i=0, \dots, \infty$), отвечающий последовательности изменения состояний упрощенного скелета с дискретным (с соответствующей частотой) течением времени; f_i^{beg} и f_i^{end} — состояния, отвечающие соответственно началу и концу жеста; S_i^{beg} и S_i^{end} — координаты концов костей упрощенного скелета для состояний f_i^{beg} и f_i^{end} соответственно (в системе координат, связанной с корневой костью упрощенного скелета); θ_i — множество изменений углов Эйлера.

Для определения траектории (θ) между жестами G_l и G_m предложено рассчитывать значения обобщенных координат следующим образом:

$$\theta_i = \theta_{i-1} + (\theta_{\text{beg}}(G_m) - \theta_{\text{end}}(G_l)) / k, \quad i = \overline{1, k-1}, \quad (8)$$

где $\theta_{\text{end}}(G_l) = \theta_0$, $\theta_{\text{end}}(G_l)$ — значения обобщенных координат для конечного состояния (f_i^{end}) жеста G_l ; $\theta_{\text{beg}}(G_m)$ — значения обобщенных координат для начального состояния (f_i^{beg}) жеста G_m ; k — количество состояний между этими двумя жестами.

Состояние между двумя жестами является фактически временем, в течение которого необходимо перейти из конечного положения одного состояния в исходное положение другого. Для автоматического определения числа состояний рассчитывается коэффициент, который является отношением числа состояний перехода между двумя жестами к расстоянию, на которое сделано перемещение некой точки на скелете,

$$\text{coef} = k^* / L, \quad (9)$$

где k^* — количество состояний перехода между двумя положениями (полученными от человека-носителя ЖР при фиксированной скорости воспроизведения жеста), L — расстояние, на которое переместилась точка на скелете.

С помощью коэффициента coef и с учетом координат (для каждого жеста) начала костей упрощенного скелета S_i^{beg} для состояния f_i^{beg} и скелета S_i^{end} для состояния f_i^{end} определяется число состояний перехода от одного жеста к другому:

$$k = \text{coef} \cdot L_{\text{max}}, \quad (10)$$

где $L_{\text{max}} = \max \left\{ \sqrt{(S_i^{\text{end}}(G_l) - S_i^{\text{beg}}(G_m))^2} \right\}$ — максимальное расстояние между двумя соответствующими точками-окончаниями каждой кости упрощенного скелета ($i=1, N-1$, N — количество костей в скелете), т.е. состояниями f_l^{end} (для жеста G_l) и f_m^{beg} (для жеста G_m).

Поскольку существуют некоторые довольно сложные жесты (воспроизводящие двумя руками сложные ассоциации из реальной жизни), возможны ситуации, когда полученная указанным образом траектория (8) приведет к коллизии при воспроизведении пространственной модели (неестественные пересечения рук, вкладывание одной руки в другую и т.д.). Для исключения таких ситуаций предлагается вычислять траектории с использованием промежуточных состоя-

ний, т.е. для таких сложных движений вводятся дополнительные промежуточные состояния, расчет траектории через которые не будет приводить к коллизиям.

При тестировании алгоритма синтеза промежуточных траекторий-переходов между жестами выявлены их реалистичность для воспроизведения жестовых конструкций. Можно констатировать, что построенные с помощью предложенного подхода траектории реалистичны и соответствуют оцифрованным траекториям-переходам.

Если разложить жест на элементы, которые однозначно его характеризуют (конфигурацию пальцев руки, локализацию выполнения, траекторию (направление) и характер движения), то становится возможным получить произвольный жест (7) не с помощью оцифровки движения носителя ЖР, а синтезированным путем. Предложенное формальное описание жеста состоит из определенных ключевых состояний обобщенного скелета (полученных оцифровкой движения носителя ЖР) и функционально рассчитанных траекторий-движений. Множества этих состояний и траекторий фиксированы. С их использованием и с помощью алгоритма синтеза промежуточных траекторий становится возможным синтезировать произвольный жест (7).

Таким образом, с помощью носителей ЖР создается фиксированное множество состояний обобщенного скелета для произвольного количества виртуальных моделей. Используя формализованное подобным образом описание произвольного жеста, автоматически синтезируется множество жестов для ЖР.

ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ЖЕСТОВОЙ РЕЧИ

Для невербального общения людей с недостатками слуха предложено создать on-line ресурс, включающий словарь основных жестов и варианты их использования в жестовых конструкциях. С помощью такого ресурса возможно решить следующие задачи; 1) расширение круга специалистов в области сурдоперевода; 2) создание автоматизированного перевода с жестового языка в режиме on-line для облегчения коммуникации глухих и слышащих; 3) унификация жестов для разных групп людей с ограниченным слухом; 4) создание новых жестов для различных областей знаний.

Для реализации on-line ресурса предложена WEB-реализация для воспроизведения трехмерного изображения в сети Интернет, использующая предложенную математическую и информационную модели для отображения ЖР с помощью виртуальной модели.

Клиентская часть on-line ресурса разработана с использованием следующих технологий:

- HTML5 — поскольку в нем реализована возможность подключения контекста WebGL;
- WebGL — для рендеринга трехмерной графики;
- JavaScript — для вызова функций WebGL и манипуляций с вершинами модели при скиннинге;
- библиотека `webgl-utils.js` — для преобразований над матрицами.

Серверная часть реализована на Microsoft ASP.NET 4.0.

Нормали, текстурные координаты, индексы вершин загружаются в видеопамети только один раз — на этапе инициализации, поскольку они не изменяются в процессе выполнения.

Параметры модели и информация по ключевым параметрам значений углов хранятся в базе данных и используются для отправки клиенту с помощью соответствующего `http-handler`.

Воспроизведенная виртуальная модель позволяет получить анимацию ЖР в реальном времени на компьютерах с минимальными требованиями: наличие видекарты и браузера, поддерживающего HTML5 и WebGL.

В тестовом приложении (рис. 3) реализован описанный механизм для построения произвольного жеста.

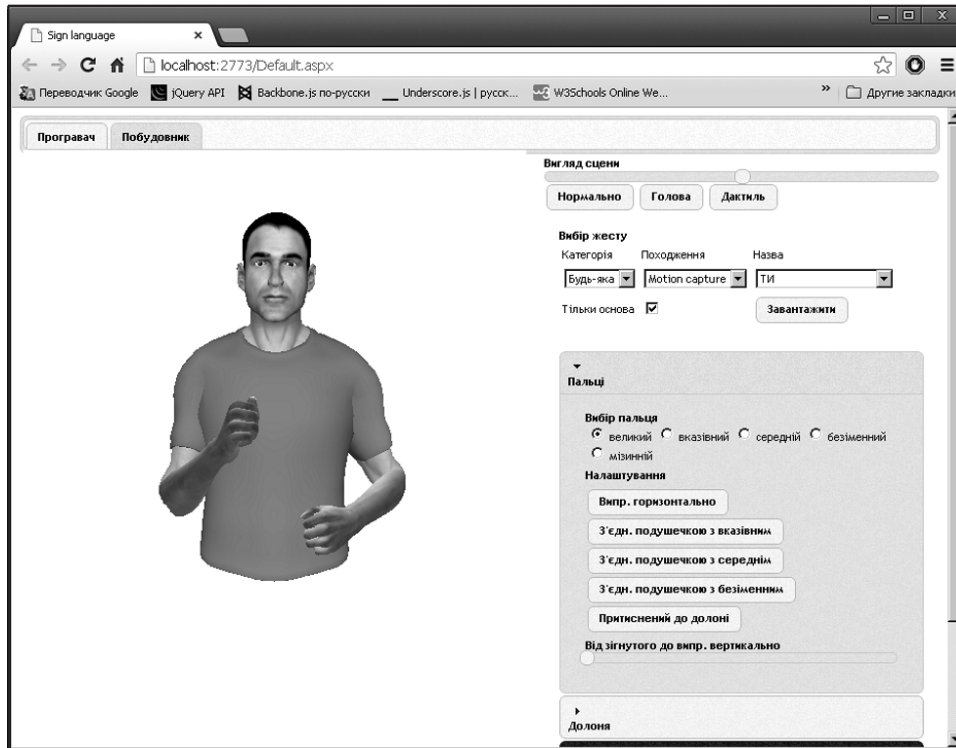


Рис. 3

ЗАКЛЮЧЕНИЕ

С помощью модели для фиксации движений ЖР и носителя ЖР были оцифрованы некоторые жесты. Используя определенные состояния этого множества жестов, были синтезированы аналогичные жесты. Сравнение и воспроизведение жестов как с оцифрованного, так и синтезированного множеств показало способность предложенной технологии достаточно реалистично воспроизводить движения на трехмерных моделях. Синтез промежуточных траекторий между любыми жестами позволил строить произвольные жестовые конструкции — предложения из множества синтезированных жестовых единиц.

Данный подход к формализации синтеза жестов позволит получать множество единиц ЖР и поддерживать их в актуальном состоянии (добавление новых единиц ЖР не требует привлечения носителей ЖР, а синтезируется как композиция устойчивых состояний и соединение устойчивых траекторий-движений).

Дальнейшие исследования направлены на совершенствование предложенной технологии для формализации описания жестов:

— нормирование необходимых множеств, а именно состояний пальцев, взаимного положения пальцев и ладони, положений ладони и направление пальцев; определение для этих множеств соответствующих значений углов положения костей обобщенного скелета;

- нормирование необходимого множества локализаций жестов и определение соответствующих значений углов положения костей обобщенного скелета;
- формализация функциональных построителей траекторий с учетом направления и характера движения;
- распознавание жестовой информации с использованием множества ключевых конфигураций.

СПИСОК ЛИТЕРАТУРЫ

1. Stokoe W.C. Sign language structure: An outline of the visual communication systems of the american deaf // *Studies in Linguistics*. — 1960. — N 8. — 78 p.
2. Система жестовой нотации SignWriting. — <http://www.signwriting.org/about/>.
3. Димскис Л.С. Изучаем жестовый язык. — М.: Академия, 2002. — 128 с.
4. Prillwitz S., Hanke T., et al. HamNoSys. Version 2.0; Hamburg Notation System for Sign Languages / An introductory guide. Hamburg: Signum, 1989. — 46 p.
5. Кривонос Ю.Г., Крак Ю.В., Бармак А.В. и др. Информационная технология для моделирования украинского языка жестов // *Искусственный интеллект*. — 2009. — № 3. — С. 186–197.
6. Кривонос Ю.Г., Крак Ю.В., Бармак О.В. Моделювання рухів віртуального персонажа для просторового відтворення жестової мови // *Реєстрація, зберігання і обробка даних*. — 2010. — 12, № 2. — С. 181–189.
7. Кириченко М.Ф., Крак Ю.В., Сорока Р.О. Оптимізація маніпуляційних роботів. — Київ: Либідь, 1990. — 144 с.
8. Бармак А.В. Информационная компьютерная технология для моделирования и управления манипуляционными и виртуальными системами // *Проблемы управления и информатики*. — 2003. — № 5. — С. 143–154.
9. Бернштейн Н.А. Очерки по физиологии движений и физиологии активности. — М.: Медицина, 1966. — 350 с.
10. Lander J. Working with motion capture file formats // *Game Developer*. — 1998. — N 1. — P. 30–37.
11. Silva V. A new interface paradigm for motion capture based animation systems. — www.visgraf.impa.br/Projects/mcapture/publ/cas97.pdf.
12. Крак Ю.В., Шкильнюк Д.В. Анализ элементов дактильного жестового языка // *Искусственный интеллект*. — 2010 — № 3. — С. 322–328.

Поступила 23.11.2012