

## **ФОРМАЛІЗАЦІЯ ЖЕСТОУТВОРЕННЯ ЗА ДОПОМОГОЮ ТРЬОХМІРНОЇ МОДЕЛІ ЛЮДИНИ ДЛЯ УКРАЇНСЬКОЇ ЖЕСТОВОЇ МОВИ**

*Ю.Г. Кривonos, Ю.В. Крак, О.В. Бармак, Г.М. Єфімов*

Інститут кібернетики ім. В.М.Глушкова НАН України,  
03680, Київ, МСП, проспект Академіка Глушкова, 40.

Запропоновано формалізацію для створення жестових одиниць за допомогою трьохмірної моделі людини для української жестової мови. Наведено формальний опис та результати тестування запропонованого підходу.

Formalization for sign units creation based on 3d human model for Ukrainian sign language is proposed. A formal description and results of testing of the proposed approach are described.

### **Вступ і актуальність проблеми**

Жестова мова (ЖМ) є одним із основних засобів спілкування для людей з вадами слуху. У ЖМ для передачі інформації використовуються візуально-кінетичні засоби - жести рук, артикуляція губ, мімічні прояви емоцій тощо, які поєднують у собі як безпосередньо жести глухих людей, так і загальноприйняті рухи (головою, тулубом тіла тощо). Українська ЖМ (УЖМ) поєднує у собі декілька різновидів жестової комунікації: дактильну абетку для передачі букв української абетки та ЖМ для передачі слів, мовних конструкцій та семантичних концептів. Система ЖМ розпадається на дві підсистеми: безпосередньо ЖМ та калькуючу ЖМ, яка повністю копіює розмовну мову. Основне, що об'єднує всі жестові системи – це спосіб передачі та сприйняття інформації. Базовою смисловою одиницею тут виступає жест (візуально-кінетичний акт), в якому беруть участь руки, міміка обличчя та артикуляція губ. Кінетична природа жесту і його візуальне сприйняття зумовлюють особливості ЖМ: можливість певним чином розташувати жест у просторі, виконати одночасно два жести двома руками тощо.

Лексичні жестові елементи і в самій ЖМ, і в дактильній абетці формуються практично однаково і можуть бути формалізовано описані наступним чином: за конфігурацією (форма руки або рук); за місцем виконання (де знаходяться руки, що виконують жест); за характером руху (як саме рухаються руки; що вони роблять).

Такий підхід вперше запропоновано У.Стокоу [1], де для опису американської ЖМ застосовано фонологічний принцип для транскрибування жестів. Під фонологічним принципом розуміється розкладання будь-якого жесту на мінімальні регулярно повторювані жестові примітиви (хіреми). Хіреми, за концепцією Стокоу, визначаються по тому, де вони жестикулюються (наприклад, у чола, у щоки, на рівні грудей тощо), і як саме вони жестикулюються (наприклад, рука рухається вперед, здійснює хвилеподібні рухи тощо). Крім того, для хірем необхідна третя ознака - форма руки або рук. Ці три ознаки (місце виконання, спосіб утворення і форма руки) використовуються у всіх сучасних системах нотації ЖМ, що призначені для формалізації елементів жестового лексикону.

Для комп'ютерної реалізації ЖМ і жестовий словник повинні бути певним чином записані, щоб комп'ютер міг обробляти і синтезувати жести. Для опису жесту за його ознаками є кілька різних систем нотації (жестової транскрипції), що дозволяють зафіксувати уявлення жесту в записі. Так, система нотації [1] базувалася на латинській системі письма, тобто для кодування хірем використовувалися букви латинського алфавіту, допоміжними елементами були арабські цифри і різні діакритики. Однак запропонований запис не підходив для освітніх цілей, оскільки був абсолютно не наочним, тому розроблялися інші системи нотацій ЖМ, що використовують символи, які більш наочно демонструють положення рук і їх переміщення. Одна з найбільш відомих систем нотації SignWriting [2], в якій знаки символізують міміку, рухи рук і тіла ЖМ і пишуться зверху вниз. До недоліків такої системи можна віднести дуже великий набір знаків (понад тисячу), наприклад, кожній специфічній формі руки відповідає окремий знак, що важко запам'ятати і використовувати на практиці.

Набагато практичніше скоротити кількість окремих випадків конфігурації руки, описуючи їх як змінні (наприклад, кількість пальців, що показуються) при одній константі (формі долоні). Такій вимозі, наприклад, задовольняє система нотації жестів, запропонована у роботі [3], де виділяється більше 30 конфігурацій рук, близько 50 характеристик місця виконання жесту і більше 70 характеристик локалізації. У нотації [3] структура жесту складається з окремих елементів і має такі сталі характеристики: 1) форму руки (рук); 2) місце розташування жесту (жестовий простір); 3) характер руху. Перший елемент припускає строго певне положення долонь і напрямки пальців, він складається з трьох частин - форми руки, положення долоні, напрямки пальців, необхідних для правильного кодування (запису) та декодування (відтворення) окремих хірем. Другий елемент структури може виконуватися над головою, біля обличчя, на рівні плечей, шиї тощо. Локалізація кожного

©Ю.Г. Кривonos, Ю.В. Крак, О.В. Бармак, Г.М. Єфімов. 2012

жесту строго стала і зміна жестового простору впливає на сенсове значення жесту в російській ЖМ. За характером виконання жести поділяються на одноручні; дворучні з однаковою формою рук; дворучні з різною формою рук. На рис. 1 показано приклад кодування жесту – числівника «чотири». Перший знак (буква П) вказує на те, що жест виконується правою рукою; другий знак іконографічно вказує форму руки – чотири розведених пальці; третій символ схематично зображує розворот долоні – вона спрямована до себе, до корпусу того, хто говорить; четвертий знак вказує на орієнтацію долоні – пальцями вгору; останній символ означає, що жест виконується на рівні грудей людини.

Найбільш широке поширення отримала система нотації HamNoSys (від англ. Hamburg Notation System) [4]. На початок 2011 року актуальна система HamNoSys версії 4.0 [5]. Вона відрізняється найбільшою пропрацьованістю і придатна для використання в комп'ютерних програмах за рахунок того, що її знаки переведені в систему кодування Unicode з відповідними комп'ютерними шрифтами. Інвентар HamNoSys дозволяє записати практично будь-який жест, що робить цю систему універсальною і придатною для будь-якої ЖМ. Фактично, за своїми функціями HamNoSys є аналогом міжнародного фонетичного алфавіту і використовується багатьма лінгвістами, що займаються дослідженням ЖМ. В системі HamNoSys спочатку задається форма кисті; потім її орієнтація за двома параметрами (напрямок пальців і розворот кисті), і місце, де розташовується рука під час жесту і характер самого руху. Жести можуть бути одноручними і дворучними, тому в другому випадку до запису додаються оператори для дворучних жестів, що дозволяють поєднати дії обох рук у рамках одного висловлювання. Конфігурація пальців і долоні, задається кількома простими знаками (див. рис. 2), що символічно зображують долоню і пальці.



Рис. 1. Приклад запису жесту – числівника «чотири» у нотації [3]

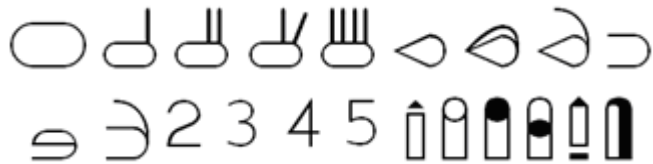


Рис. 2. Символи HamNoSys, що призначені для формального запису конфігурації пальців та долоні

Одним з найбільш ефективних засобів навчання і людино-машинної взаємодії є мультимедійні комп'ютерні програми, тому, створення інформаційних програм, здатних працювати з ЖМ (синтезувати і розпізнавати), є одним з пріоритетних завдань при роботі з глухими і слабочуючими людьми. Можливим варіантом комп'ютерного синтезу ЖМ є використання анімованого аватара людини, який керується за допомогою символів жестової нотації, описуючи необхідні конфігурації рук і різні типи рухів. Жести з лексику в такій системі синтезу є ланцюжком символів в обраній нотації, тому словник може легко модифікуватися і розширюватися без використання спеціального обладнання.

### Мета роботи та постановка задачі

Для УЖМ запропонована концепція [6] інформаційної технології невербального спілкування людей з вадами слуху. Концепція полягає у використанні мультимедійних можливостей сучасної комп'ютерної техніки та систем комунікацій а саме - використання просторових моделей людей для відтворення ЖМ. В рамках цієї концепції запропонована математична модель та її комп'ютерна реалізація для отримання та відтворення ЖМ за допомогою трьохмірних моделей людей. Для реалізації можливостей генерації анімації процесу мовлення за допомогою ЖМ з використанням віртуальних моделей людей були розроблені відповідні інформаційні та математичні моделі [7, 8]:

- 1) інформаційна та математична моделі для фіксації морфем (мінімально значимих одиниць) жестової мови;
- 2) технологія та відповідне програмне забезпечення для отримання, збереження та відтворення жестів.

Запропонована технологія моделювання рухів віртуального персонажа для просторового відтворення жестової мови крім явних переваг перед іншими системами містить суттєвий недолік: необхідність оцифрування всієї множини жестів за допомогою носія ЖМ для кожної нової моделі. Затратність такого підходу як по часу так і по вартості спонукає до пошуку інших підходів, що враховують таке обмеження. Крім того існування систем нотацій для ЖМ [2–5] спонукає до пошуку такої моделі формалізації ЖМ, яка б дозволила описувати жести у цих системах. Метою даної роботи є розробка такої моделі формалізації для українського ЖМ.

Отже, сформульована наступна *постановка задачі*: запропонувати модель для формалізації жестових одиниць з метою створення побудовника довільного жесту на віртуальній моделі з використанням фіксованих множин ключових станів, траєкторій тощо.

### Формальний опис фіксації жесту

Процес відтворення жесту трьохмірною моделлю людини можна вважати анімацією з відповідною частотою різних станів скелетної моделі людини. Скелетна модель людини спрощено відтворює скелет живої людини. Її можна формалізувати як ієрархічну структуру яка складається з поєднаних кінематичних пар, які відтворюють основні кістки скелету людини. Сучасні пакети трьохмірного моделювання (Poser, 3D Studio Max)

вміють генерувати анімацію з допомогою віртуальної статичної моделі та інформації про зміну відповідних кутів скелета. Отже, для формального опису процесу фіксації жесту можна використовувати множину, що відображає спрощений скелет людини та зміни значення кутів Ейлера і порядок їх застосування для відповідних кісток цього скелету з плином часу (дискретно, з відповідною частотою (1/30 сек тощо) [7, 8]:

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

де  $H_i$  –  $i$ -а кістка скелету ( $i=0, \dots, N-1$ ,  $N$  – кількість кісток у скелеті);  $k$  – індекс кістки-предка;  $d_i = [x_i, y_i, z_i]^T$  – координати точки – кінця кістки у системі координат, яка пов'язана з початком цієї кістки;

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

де  $M_i$  – для  $i$ -ї кістки значення кутів Ейлера та порядок застосувань обертань для кістки з плином часу;  $order_i \in \{1, \dots, 6\}$  – порядок застосування обертання навколо відповідних координатних осей для  $i$ -ї кістки (1-XYZ, 2-XZY, 3-YXZ, 4-YZX, 5-ZXY, 6-ZYX);  $\theta = (\theta^j)$ ,  $\theta^j = \{\theta_i^j : \theta_i^j = (\varphi_{i,x}^j, \varphi_{i,y}^j, \varphi_{i,z}^j)\}$  – множина зміни кутів Ейлера для  $i$ -ї кістки з плинним часу ( $j = 0, \dots, K-1$ ,  $K$  – кількість кадрів відтворення руху з заданою частотою).

Для отримання множини кутів, що відтворюють процес анімації жестів використовувалась технологія «захоплення» руху (Motion Capture). Множина кутів отримувалась з людини – носія ЖМ. Це дозволило отримати досить реалістичні, людиноподібні рухи.

Для синтезу процесу анімації жестів та міміки трьохмірною моделлю людини було запропоновано наступний формальний опис, який використовує відповідні множини параметрів та алгоритми роботи з ними. Трьохмірна модель людини, на якій реалізується процес анімації жестів та міміки, має наступні атрибути:  $V = \{v_i : v_i = \{x, y, z\}\}$  – множина вершин трикутників для триангуляції поверхні трьохмірної моделі людини;  $N = \{n_i : n_i = \{x, y, z\}\}$  – множина нормалей до вершин;  $T = \{t_i : t_i = \{t, s\}\}$  – множина текстурних координат до вершин;  $V^{ind} = \{V_i^{ind} : V_i^{ind} = \{k_1, k_2, k_3\}\}$  – множина індексів, які вказують порядок побудови трикутників з множини вершин;  $I = \{I_i : I_i = \{img\}\}$  – множина фотографічних зображень елементів моделі – текстури.

Для моделювання скелетної анімації потрібно вміти розраховувати нові значення вершин трикутників ( $V$ ). Для цього запропоновано використовувати механізм скінінга. Скінінг можна означити як алгоритм прив'язки множини вершин трикутників, які означають поверхню моделі до значень кутів скелету. Тоді модель скелетної анімації можна формалізувати наступним чином:  $MH = \{MH_i : MH_i = \{k, \{l_1, \dots, l_m\}, d_i, Glb, Order_i\}\}$  – опис спрощеного скелету людини (ієрархія кісток) для реалізації скелетної анімації, де  $MH_i$  –  $i$ -а кістка скелету ( $i = 0, \dots, N-1$ ,  $N$  – кількість кісток у скелеті);  $k$  – індекс кістки-предка;  $\{l_1, \dots, l_m\}$  – множина індексів нащадків кістки;  $d_i = [x_i, y_i, z_i]^T$  – координати точки – кінця кістки у системі координат, яка пов'язана з початком цієї кістки;  $Glb$  – вектор для визначення координат кістки у глобальній системі координат,  $Order_i$  – порядок застосування обертання.  $Skin = \{Skin_i : Skin_i = \{(IndexVertex_1, Weight_1), \dots\}\}$  – множини вершин, які впливають на поточну вершину при зміні кутів.

Скінінг розраховується для кожної вершини наступним чином:

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

де  $n$  – кількість кісток, пов'язаних з вершиною  $v$ ;  $IBM_{H_i}$  – інверсна матриця для кістки  $H_i$ ;  $JM_{H_i}$  – матриця переміщення для кістки  $H_i$ ;  $JW_{H_i}$  – ваговий коефіцієнт для застосування впливу точок кістки  $H_i$  на вершину  $v$ .

Для моделювання анімації промовляння та емоційного забарвлення запропоновано використовувати механізм морфінгу. Морфінг можна означити як алгоритм плавного переходу від одного стану об'єкта до іншого. При використанні морфінгу використовуються лише опорні стани, з допомогою яких розраховуються проміжні стани і моделюється процес анімації. Модель морфінга міміки промовляння можна формалізувати наступним чином:

Відображення або побудова міміки на обличчі трьохмірної моделі досягається при застосуванні відносного (сегментного) морфінгу до моделі голови і мімічного процесу. Формула відносного морфінгу для  $M$  морфів у формалізмі моделі матиме наступний вигляд:

$$V' = V + \sum_{m=1}^M w_m \cdot TM_m,$$

де  $w_m$  – вагові коефіцієнти,  $V$  – меш базової моделі без морфінгу,  $TM_m$  – вхідний меш (морф) для блендінгу (blending). Результатом операції є лінійна комбінація мешів моделі та мімік.

Метод сегментного морфінгу дає змогу: формувати декілька виразів обличчя на основі невеликої кількості морфів та змінювати стан обличчя при анімації промовляння. Додатковою перевагою сегментного морфінгу є те, що можна анімувати щелепу незалежно від губ та очі (кліпання), незалежно від емоційного прояву на обличчі.

### Модель відтворення жесту

Для відтворення процесу анімації жестів та міміки трьохмірною моделлю людини було створено засіб, який реалізує скелетну (для відтворення жесту) та морфемну (для відтворення міміки промовляння та емоцій) анімацію. Реалізована відповідна програмна функціональність, що відтворює по означеним атрибутам модель людини (рис. 3) та, з використанням алгоритмів скінінга та морфінгу відтворює анімацію жестових одиниць (рис. 4) [7, 8].



Рис. 3. Відтворення трьохмірної моделі людини



Рис. 4. Кадри анімації жесту «Ти»

Для формалізації жесту використовується підхід [1], згідно якого, запропоновано розкласти жест на наступні елементи, які його однозначно характеризуватимуть: 1) конфігурацію пальців руки; 2) локалізацію виконання; 3) траєкторію (напрямок) та характер руху.

Можливі у жестовій мові основні конфігурації пальців руки подаються за допомогою наступних параметрів (стан кожного пальця, взаємне положення пальців та долоні, положення долоні, напрямок пальців):






- для кожного з пальців (великого, вказівного, середнього, безіменного та мізинця) введені наступні можливі стани: випрямлений, випрямлений в горизонтальному положенні, зігнутий, напівзігнутий, притиснений до долоні або інших пальців (табл. 1);

Таблиця 1. Можливі стани пальців руки

Можливі стани		Пальці руки				
		Великий	Вказівний	Середній	Безіменний	Мізинець
1.	Випрямлений					
2.	Випрямлений в горизонтальному положенні					
3.	Зігнутий					
4.	Напівзігнутий					
5.	Притиснутий до долоні чи інших пальців					

- для подання взаємного положення пальців та долоні введені наступні ознаки: пальці розведені або зведені (стиснуті); лежать один на одному; притиснуті подушечками; притиснуті один поверх іншого (інших) (див. табл. 2);

Таблиця 2. Взаємне положення пальців та долоні:

Можливі положення		Пальці руки				
		Великий	Вказівний	Середній	Безіменний	Мізинець
1.	Розведені / Зведені (стиснуті)					
2.	Лежать один на одному					
3.	Притиснені подушечками					
4.	Притиснені один поверх іншого (інших)					

- положення долоні визначені наступними значеннями: вниз; вверх; до себе; від себе; вліво; вправо; одна до однієї горизонтально; одна долоня вверх, інша вниз, одна над іншою; одна до іншої вертикально, торкаються подушечками пальців; вверх, торкаючись одна до іншої тильною та внутрішньою сторонами; дві долоні до себе не торкаючись одна іншою; донна до іншою з торканням; одна вверх, інша вниз, торкаючись одна однієї; обидві вниз з торканням одна однієї сторонами вказівних пальців; обидві від себе з торканням одна однієї сторонами вказівних пальців;
- напрямок пальців визначений наступними значеннями: вниз; вверх; від себе; до себе; вліво; вправо; горизонтально, назустріч з торканням; горизонтально, назустріч без торкання; вертикально, назустріч без торкання; вертикально, назустріч з торканням; горизонтально, в протилежні сторони з торканням; горизонтально, в протилежні сторони без торкання.

Для визначення локалізації (місця) виконання жесту запропоновано використовувати наступні горизонтальні рівні: над маківкою; на рівні лобу; лінія брів; лінія очей, рівень правого і лівого вуха; рівень щік; рот; підборіддя; під підборіддям; лінія шиї; плечі; груди; талія; лінія нижче талії. Для кожного горизонтального рівня вказується вертикаль: зліва, в центрі, справа. Для кожної області (перетин горизонталі та вертикалі) вказується ознака: з дотиком чи без дотику.

Напрямок та характер руху вказується за допомогою наступних характеристик: вліво, вправо, вверх, вниз, під кутом вліво вверх, під кутом вправо вверх, під кутом вправо вниз, під кутом вліво вниз, назустріч горизонтально, в протилежні сторони – горизонтально, назустріч вертикально, в різні сторони – вертикально, догори-донизу – за чергою, до себе, від себе. Для кожної перерахованої характеристики вказуються параметри: рухи повторюються чи ні; швидкість руху: звичайна, швидко, повільно (плавно), уривчато. Крім того можливі рухи по певним траєкторіям: півколом (вниз чи догори, від себе чи до себе), колом (за часовою стрілкою чи проти), ромбом, хвилеподібно (дрібна чи глибока хвиля), криволінійно, східчастий рух тощо.

У роботах [7, 8] авторами запропонована технологія для синтезу проміжних траєкторій-переходів між жестовими одиницями у конструкціях-реченнях. В рамках моделі, що використовується, знайдено функціональний перехід від кінцевого стану спрощеного скелету (вектор узагальнених координат – значення кутів обертання навколо відповідних осей при кожній ланці (кістці)) одного жесту до початкового стану (інший вектор узагальнених координат) іншого жесту. Даний функціональний перехід породжував траєкторію, що візуально близька до траєкторії людини при аналогічному переході.

Для формалізації жесту та з метою подальшого використання механізму аглютинації (для створення жестових конструкцій типу речень) використано наступну множину:

$$G = \{G_i : G_i = \{f_i^{beg}, f_i^{end}, S_i^{beg}, S_i^{end}, \theta_i\}\}, \quad (3)$$

де  $G_i$  – конкретний жест який відповідає послідовності зміни станів спрощеного скелета з дискретним (з відповідною частотою) плином часу (1)(2);  $f_i^{beg}$  – номер стану, що відповідає початку жесту;  $f_i^{end}$  – номер стану, що відповідає кінцю жесту;  $S_i^{beg} = \{(x, y, z)\}$  – координати кінців кісток спрощеного скелета для стану  $f_i^{beg}$  (у системі координат, пов'язаній з кореневою кісткою спрощеного скелета);  $S_i^{end}$  – аналогічно для стану  $f_i^{end}$ ;  $\theta_i$  – множина змін кутів Ейлера (із (2)). Для розрахунку траєкторії ( $\theta$ ) між жестами  $G_i$  та  $G_m$  значення узагальнених координат знаходяться наступним чином:

$$\begin{aligned} \theta_0 &= \theta_{end}(G_i), \\ \theta_i &= \theta_{i-1} + (\theta_{beg}(G_m) - \theta_{end}(G_i)) / k, i = 1, \dots, k - 1 \end{aligned} \quad (4)$$

де  $\theta_{end}(G_i)$  – значення узагальнених координат для останнього стану ( $f_i^{end}$ ) жесту  $G_i$ ;  $\theta_{beg}(G_m)$  – значення узагальнених координат для першого стану ( $f_m^{beg}$ ) жесту  $G_m$ ;  $k$  – кількість станів між цими двома жестами.

Кількість станів між двома жестами є фактично часом за який потрібно перейти з кінцевого положення одного стану до початкового положення іншого. Для автоматичного визначення кількості станів обраховується коефіцієнт який є відношенням кількості станів переходу між двома жестами до відстані, на яку зроблено переміщення деякої точки скелету, тобто:  $coef = \frac{k}{L}$ , де  $k$  – кількість станів між двома положеннями (отриманими з людини-носія жестової мови при фіксованій швидкості показу жесту);  $L$  – відстань на яку перемістилась деяка точка скелету.

Оцифрування жестів множини  $G$  проводилось у одному (середньому) темпі. Розрахувавши коефіцієнт  $coef$  для різних жестів, отримується його середнє значення яке використовується для розрахунку кількості перехідних станів ( $K$ ) між двома жестами.

За коефіцієнтом  $coef$ , та координатами (для кожного жесту) початку кісток спрощеного скелета  $S_i^{beg}$  для стану  $f_i^{beg}$  та  $S_i^{end}$  для стану  $f_i^{end}$  визначається кількість станів переходу від одного жесту до іншого ( $k$ ):  $k = coef \cdot L_{max}$ , де  $L_{max} = \max\left\{\sqrt{(S_i^{end}(G_i) - S_i^{beg}(G_m))^2}\right\}$  – максимальна відстань між двома відповідними точками-кінцями кожної кістки спрощеного скелету ( $i = 0, \dots, N-1$ ,  $N$  – кількість кісток у скелеті) між станами  $f_i^{end}$  (для жесту  $G_i$ ) і  $f_m^{beg}$  (для жесту  $G_m$ ).

Враховуючи те, що існують деякі доволі складні жести (що відтворюють двома руками складні асоціації з реального життя), можливі ситуації коли отримана означеним чином траєкторія (4) приведе до колізії при відтворенні просторовою моделлю (неприродні перетини рук, занурення однієї руки в іншу тощо). Для подолання таких ситуацій пропонується обраховувати траєкторії з використанням проміжних станів. Тобто для таких складних рухів вводяться додаткові проміжні стани, розрахунок траєкторії через які, не приводитиме до колізій.

### Аналіз отриманих результатів

Тестування алгоритму синтезу проміжних траєкторій-переходів між жестами показало їх спроможність та реалістичність для відтворення, побудованих на основі жестів, жестових конструкцій. На рис. 5 та 6 показано побудовані та реальні траєкторії для проміжних станів між жестами у жестових конструкціях. Можна констатувати, що побудовані, за допомогою запропонованого підходу, траєкторії реалістичні і відповідають оцифрованим траєкторіям-переходам.

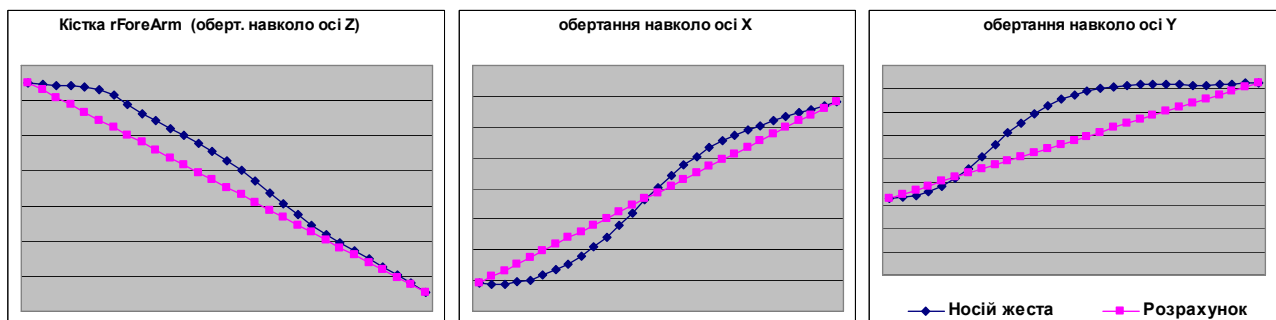


Рис. 5. Порівняння траєкторій руху при переході між жестами «ТИ» і «Початковий»



Рис. 6. Кадри переходу від жесту «ТИ» до початкового жесту

Якщо розкласти жест на елементи, що його однозначно характеризують (конфігурацію пальців руки; локалізацію виконання; траєкторію (напрямок) та характер руху), то стає можливим отримати довільний жест з



множини (3) не за допомогою оцифрування руху носія ЖМ, а синтетичним шляхом. Запропонований формальний опис жесту складається з певних ключових станів узагальненого скелета (отриманих оцифруванням руху з носія ЖМ) та функціонально розрахованих траєкторій-рухів. Множини цих станів та траєкторій фіксовані, тому за допомогою їх, та алгоритму синтезу проміжних траєкторій стає можливим синтезувати довільний жест.

Таким чином, за допомогою носіїв ЖМ створюється фіксована множина станів узагальненого скелета для довільної кількості віртуальних моделей. Використовуючи формалізований опис довільного жесту автоматично синтезується множина жестів для ЖМ.

В рамках запропонованої авторами концепції [7, 8] інформаційної технології невербального спілкування людей з вадами слуху запропоновано створити on-line ресурс, що включатиме у себе жестівник основних жестів та використання їх у жестових конструкціях. За допомогою такого ресурсу планується вирішити наступні задачі: 1) розширення кола фахівців в області сурдоперекладу; 2) створення автоматизованого перекладу жестової мови в режимі on-line для полегшення комунікації глухих та чуючих; 3) уніфікації жестів для різних груп людей з вадами слуху; 4) створення нових жестів для різних областей знань.

Для реалізації on-line ресурсу, запропонована WEB-реалізація для відтворення трьохмірного зображення у мережі Інтернет, що використовує запропоновану математичну та інформаційну моделі для відображення ЖМ за допомогою віртуальної моделі.

Клієнтська частина on-line ресурсу розроблена за допомогою наступних технологій:

- HTML5 – тому, що в ньому реалізований елемент «canvas», що дозволяє підключати контекст WebGL;
- WebGL – для рендерингу тривимірної графіки;
- JavaScript – для виклику функцій WebGL та маніпуляцій з вершинами моделі при скіннігу;
- Бібліотека «webgl-utils.js» - для перетворень над матрицями.

Серверна частина реалізована на Microsoft ASP.NET 4.0.

Рендерінг віртуальної моделі реалізовано за допомогою WebGL з використанням тег'а «canvas» HTML5: `<canvas id="webgl-canvas" style="border: none;" width="800" height="600"></canvas>`. Прив'язка контексту WebGL виконується у методі `initGL: gl = canvas.getContext("experimental-webgl")`.

Нормалі, текстурні координати, індекси вершин - завантажуються до відео-пам'яті лише один раз – на етапі ініціалізації, оскільки вони не змінюються під час виконання:

```
...
cubeVertexNormalBuffer = gl.createBuffer();
gl.bindBuffer(gl.ARRAY_BUFFER, cubeVertexNormalBuffer);
gl.bufferData(gl.ARRAY_BUFFER, new window.Float32Array(vertexNormals), gl.STATIC_DRAW);
cubeVertexNormalBuffer.itemSize = 3;
cubeVertexNormalBuffer.numItems = vertexNormals.length;
...
```

Так як, до вершин застосовується скіннінг, то вони розраховуються під час рендерингу кожного кадру:

```
...
Skin(0, matIdent);

cubeVertexPositionBuffer = gl.createBuffer();
gl.bindBuffer(gl.ARRAY_BUFFER, cubeVertexPositionBuffer);
gl.bufferData(gl.ARRAY_BUFFER, new window.Float32Array(points), gl.STATIC_DRAW);
cubeVertexPositionBuffer.itemSize = 3;
cubeVertexPositionBuffer.numItems = points.length;
...
```

Функція скіннігу реалізована рекурсивно і проходить по всіх вузлах моделі починаючи з кореневого.

```
function Skin(nodeIndex, mat) {
    var m = mat4.multiply(mat, transforms[frameNum][nodeIndex]);

    var i;
    var v;
    var bone;
    var weight;

    for(i = 0; i < bones[nodeIndex].length; i++) {
        bone = bones[nodeIndex][i][0];
        weight = bones[nodeIndex][i][1];
        v = [vertices[bone * 3], vertices[bone * 3 + 1], vertices[bone * 3 + 2]];

        points[bone * 3] += (v[0] * m[0] + v[1] * m[4] + v[2] * m[8] + m[12]) * weight;
        points[bone * 3 + 1] += (v[0] * m[1] + v[1] * m[5] + v[2] * m[9] + m[13]) * weight;
        points[bone * 3 + 2] += (v[0] * m[2] + v[1] * m[6] + v[2] * m[10] + m[14]) * weight;
    }

    for (i = 0; i < children[nodeIndex].length; i++) {
        Skin(children[nodeIndex][i], m);
    }
}
```

Параметри моделі та інформація по ключових параметрах значень кутів зберігаються у базі даних та використовуються для відправки його клієнту за допомогою відповідного http-handler'a.

Запропонована реалізація відтворення віртуальної моделі дозволяє отримати анімацію ЖМ у реальному часі на комп'ютерах з мінімальними вимогами: наявність відео-карти і браузера, що підтримує HTML5 і WebGL (рис. 7).



Рис. 7. Рендерінг віртуальної моделі у браузері (HTML5, WebGL)

## **Висновки**

Використовуючи модель для фіксації рухів, що відтворюють УЖМ та носія ЖМ, була оцифрована множина із 50 жестів. Використовуючи певні стани із цієї множини були синтезовані аналогічні жести. Порівняння та відтворення жестів як із оцифрованої так і з синтезованої множин показало спроможність запропонованої технології досить реалістично відтворювати рухи на трьохмірній моделі. Синтез проміжних траєкторій між довільними жєстами дозволив будувати довільні жєстові конструкції – речення з множини синтезованих жєстових одиниць.

Розроблене, на основі запропонованої технології, програмне забезпечення може стати основою для створення стандарту ЖМ. Це дозволить вирішити проблему відмінностей для одних і тих самих жестів, яка виникає від того, що жести, які вивчаються, містять особливості конкретного викладача. Запропонований підхід до формалізації жєстоутворення дозволить отримувати множину одиниць ЖМ та підтримувати її у актуальному стані (додавання нових одиниць ЖМ не потребуватиме залучення носіїв ЖМ а буде синтезуватись як композиція сталих станів та синтез сталих траєкторій-рухів).

Подальші дослідження направленні на вдосконалення запропонованої технології для формалізації жєстоутворення для УЖМ: унормування необхідних множин: станів пальців, взаємного положення пальців та долоні, положень долоні та напрямків пальців; визначення для цих множин відповідних значень кутів для станів кісток узагальненого скелета; унормування необхідної множини локалізацій жестів та визначення відповідних значень кутів для кісток узагальненого скелета; формалізація функціональних побудовників траєкторій з урахуванням напрямку та характеру руху.

1. *Stokoe W.C.* Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf. // *Studies in Linguistics: Occasional papers.* – 1960. – № 8. – 78 p.
2. Система жєстової нотації SignWriting // Доступ: <http://www.signwriting.org/about/>.
3. *Димскис Л.С.* Изучаем жєстовый 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. Система жєстової нотації Hamburg Notation System // Доступ: <http://www.sign-lang.uni-hamburg.de/projects/hamnosys.html>.
6. *Кривонос Ю.Г.* [та ін.]. Інформаційна технологія для моделювання української мови жестів // *Штучний інтелект.* – 2009. – № 3. – С. 186–197.
7. *Кривонос Ю.Г., Крак Ю.В., Бармак О.В.* Моделювання рухів віртуального персонажа для просторового відтворення жєстової мови // *Реєстрація, зберігання і обробка даних.* – 2010. – Т. 12, № 2. – С. 181–189.
8. *Кривонос Ю.Г., Крак Ю.В., Бармак О.В.* Комп'ютерне моделювання української жєстової мови. // *Системні дослідження та інформаційні технології.* – 2011. – № 4. – С. 48–60.