

**А. В. Лукоянычев**, аспирант, dizzystyle@yandex.com,  
Новосибирский государственный технический университет

# Интегрированная система обучения жестовому языку

*Рассмотрена разработанная автором интегрированная система обучения жестовому языку, ориентированная на широкий круг пользователей. В системе используются нотационная запись Димскис для реализации отображения жестового языка и межплатформенная среда разработки Unity 3D для управления анимированным персонажем. Предложен подход для описания движений 3D-персонажа на основе метаязыка. Рассмотрена методика создания справочника жестового языка, приведены преимущества такого подхода. Интегрированная система, кроме полноценного инструментария для создания справочника, представляет возможность проверки знаний пользователя для мобильных и стационарных устройств. Рассматриваются этапы реализации данного режима. Приведены результаты экспериментального исследования.*

**Ключевые слова:** жестовый язык, нотация Димскис, Unity 3D, метаязык управления аватаром, локализация жеста, идентификация жеста

## Введение

Обучение жестовому языку необходимо людям, контактирующим со слабослышащими. Создание обучающих жестовому языку систем актуально и для решения задачи повышения уровня автоматизации и роботизации всех сфер деятельности человека. В настоящее время уделяется значительное внимание технологиям организации эффективного и более естественного человеко-машинного взаимодействия. С помощью жестов рук можно передавать управляющие команды интеллектуальной информационной системе на расстоянии, в зашумленной обстановке, в экстремальных условиях, когда речевые команды затруднительны.

Классические пользовательские интерфейсы уступают место естественным для человека интерфейсам. Интеллектуальные информационные системы начинают применять в медицине, образовании, социальной сфере, военном деле, а также при управлении робототехническими комплексами [1–4].

В документе "Стратегия развития отрасли информационных технологий в Российской Федерации на 2014–2020 годы и на перспективу до 2025 года" выделены приоритетные направления государственной политики по исследованиям и разработкам в области информационных технологий. Один из пунктов этого документа формулирует следующее направление: "Новые человеко-машинные интерфейсы, включая новые

методы использования жестов, ... а также новые программные средства и устройства, повышающие социальную адаптацию людей с ограниченными возможностями". В соответствии с этим положением необходима разработка технологий и средств автоматизированного обучения языку жестов глухих людей, а также людей, использующих жесты для эффективного взаимодействия со слабослышащими, для организации межчеловеческого и человеко-машинного взаимодействия.

Несмотря на большое число работ в данном направлении, вопросам демонстрации и обучения жестовому языку (ЖЯ) в настоящее время уделяется недостаточное внимание. На настоящее время не существует действующей, полностью автоматизированной универсальной расширяемой системы обучения и проверки знаний ЖЯ.

Создание таких систем связано с рядом проблемных вопросов: сложность использования нотационных записей для представления ЖЯ; ограниченный круг разработчиков систем, что не позволяет сторонним заинтересованным лицам пополнять справочник ЖЯ; ориентация систем на специальное оборудование для описания и распознавания жестов; стремление решать общие проблемы, связанные с сурдопереводом, даже в рамках обучающей системы. В результате возможности таких систем обычно ограничиваются небольшим набором используемых жестов или имеются существенные ограничения на минимальную сложность оборудования, что сужает круг пользователей этих систем [5, 6].

## Нотации жестового языка

Анализ возможности применения компьютеров для отображения ЖЯ показывает, что наиболее перспективным является использование 3D-анимации для демонстрации жестов.

Основой любого ЖЯ является жестовая транскрипция (нотационная запись), на основе которой составляется справочник ЖЯ. Существует несколько систем нотаций, различных по возможностям и принципам представления жеста в записи.

В настоящее время международными исследователями ЖЯ используются следующие системы: нотация Стокоу; система транскрипции Berkeley; система транскрипции ЖЯ HamNoSys; система нотации SignWriting. Кроме перечисленных систем существуют и другие менее известные системы записи ASL, к числу которых относятся Sign Script, Si5s, ASL-phabet. У каждой из них есть свои преимущества и недостатки [6—8]. Для учета особенностей русского жестового языка (РЖЯ) разработана система нотации Л. Димскис [9].

Нотации Димскис содержат около 150 нотационных записей. Конфигурация пальцев (форма руки) представляется в виде иконографического изображения и содержит менее 40 вариантов. Место исполнения жеста описывается 50 нотационными записями, а характеристик локализации жеста — менее 80.

Иконографическое представление конфигурации пальцев на современном уровне развития 3D-графических редакторов позволяет достаточно легко создавать такие изображения на компьютере. Такое изображение формы руки уменьшает число символов и повышает наглядность при описании жеста. Ограниченный интуитивно понятный базовый набор символов в нотации облегчает создание описания транскрипционных записей даже сложных движений пальцев и рук анимированного персонажа. В нотации Димскис отсутствуют взаимодействия с окружающей обстановкой и описания не мануальных действий (глаза, рот). Это обстоятельство свидетельствует о том, что система нотации Димскис, как и любая другая нотационная компьютерная система, должна совершенствоваться и развиваться. Так, в системе SignWriting на настоящее время насчитывается более 38 тыс. знаков [5]. Массовое использование нотации Димскис в РФ потенциально расширяет круг пользователей и составителей справочника и уроков [10]. Применение нотации Димскис для создания открытой интегрированной системы обучения ЖЯ на текущий момент является приемлемым обоснованным решением.

Для управления анимированным 3D-персонажем (аватаром) существует несколько графических пакетов (игровых движков). Наиболее универсальными и развитыми являются Unity 3D, Cry ENGINE и UDK [11]. Сравнение их характеристик показало, что наиболее подходящим для решения поставленных задач является пакет Unity 3D. Он позволяет создавать приложения под несколько операционных систем для персональных компьютеров и мобильных устройств. Возможно использование этого приложения в виде интернет-сервиса.

## Методика создания мультимедийного справочника жестового языка

В нотации Димскис все слова и дактилемы начинаются с пятисимвольного описания положения руки и пальцев. Эти символы описывают: характер исполнения жеста (одноручный, двуручный); положение пальцев; направление ладони; направление пальцев; место расположения жеста. Затем следуют дополнительные символы, с помощью которых отображается динамика выполнения жеста.

Анализ анатомического строения руки человека показывает, что контроль состояния движения руки можно осуществлять посредством управления 19 подвижными элементами для одной руки.

Для удобства работы с Unity 3D разработаны метаязык (скриптовый язык) и программное обеспечение, которое позволяет переводить скриптовый текстовый формат в анимационное движение в Unity 3D. Создание метаязыка позволило упростить управление процессом отображения движений аватара. Это обстоятельство позволяет наглядно и оперативно корректировать файл. При этом сокращается объем файлов клипов, что особенно актуально для мобильной реализации.

Нотация Димскис не является полностью компьютерно-ориентированной, поэтому в рамках метаязыка расширена нотационная запись Димскис. Введены дополнительные и служебные символы, управляющие слова, специальные координаты [12]. Разработаны специальные управляющие символы и правила для синхронизации процессов, позволяющие реализовать движение нескольких элементов за разное время (с разной скоростью), создавать сложные движения, осуществлять возврат в начальное состояние, отменять перемещения отдельных подвижных элементов за счет использования "фиктивного" движения, изменять только отдельно взятые координаты руки аватара. Данная модификация нотации Димскис позволяет формировать набор элементарных анимаций

(клипов), которые составляют основу визуализированных нотаций, и создавать эффект непрерывности демонстрации жестов, реализовывать любое желаемое движение. Правила, разработанные для синхронизации процессов, отличаются наглядностью представления, уменьшением сложности составления описания и производительностью по сравнению со стандартными методами анимации. Структура разработанной интегрированной системы обучения РЖЯ представлена на рис. 1.

Инструментарий интегрированной системы ориентирован на разработчиков и на пользователей и, соответственно, имеет два режима: "Разработчик" и "Пользователь".

Технологический инструментарий для разработчика включает все необходимое программное обеспечение для создания анимационного файла и демонстрации жеста, для соединения созданного клипа с нотацией Димскис. На основе созданных нотаций разрабатываются мультимедийный справочник по РЖЯ и уроки по обучению РЖЯ. Редакторы справочника и уроков в режиме "Разработчик" доступны в сети Интернет и могут использоваться сторонними специалистами в области ЖЯ. Разработанная методика создания справочника и уроков РЖЯ позволяет оперативно

проверять и настраивать сценарии демонстрации слов, высказываний и уроков на их основе.

Просмотр библиотеки дактилем и слов справочника и разработанных уроков возможен сторонними пользователями как через сеть Интернет, так и на персональных и мобильных устройствах при загрузке модулей режима "Пользователь".

Основным модулем для разработки анимационных файлов (клипов) является "Редактор клипов". В редакторе доступны для отображения все созданные анимационные файлы. Разработчик имеет возможность визуально создавать клипы и сценарии любой сложности за счет полноценного инструментария, используя всю иерархию библиотеки от простейших анимаций до составных (рис. 2, а). В результате создается файл клипа на метаязыке, который можно редактировать либо во встроенном, либо в любом текстовом редакторе (см. рис. 2, б, в).

Следующим шагом является связывание созданных клипов с нотациями Димскис. "Редактор знаков нотации" позволяет создать словесное описание выбранного анимированного символа и указать его расположение в библиотеке знаков нотации. На рис. 3 представлен пример связки анимационного клипа с элементом нотации Димскис.

На основе полученных анимированных знаков в "Редакторе справочника" в соответствии с нотацией Димскис составляются описания дактилем и слов, сценарии реализации которых вносятся в справочник (рис. 4). Сформированные записи в справочнике РЖЯ в дальнейшем могут использоваться для составления новых слов и уроков.

Справочник РЖЯ включает в себя свыше 3000 слов, поэтому редакторы справочника и уроков в режиме "Разработчик" могут использоваться сторонними специалистами в области ЖЯ, а интегрированная система доступна через сеть Интернет. Такой подход позволяет привлечь заинтересованных лиц к наполнению справочника и уроков. Для изучения РЖЯ в режиме "Пользователь" возможна загрузка на персональные компьютеры и мобильные устройства модулей просмотра библиотеки дактилем и слов, а также созданных уроков.

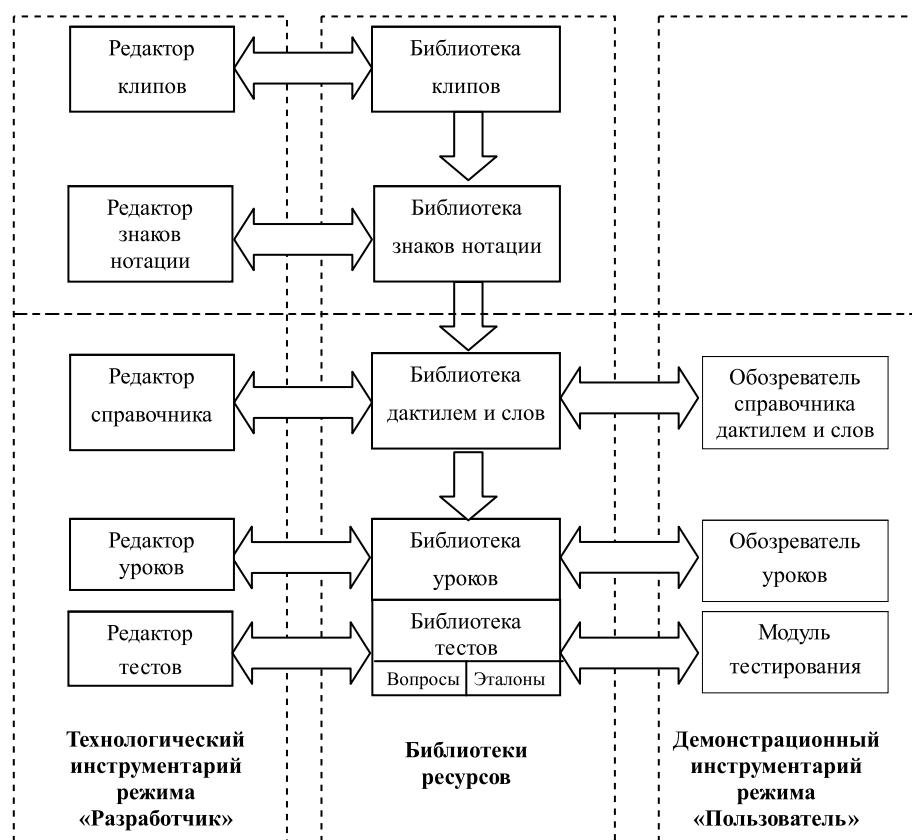


Рис. 1. Структура интегрированной системы обучения РЖЯ

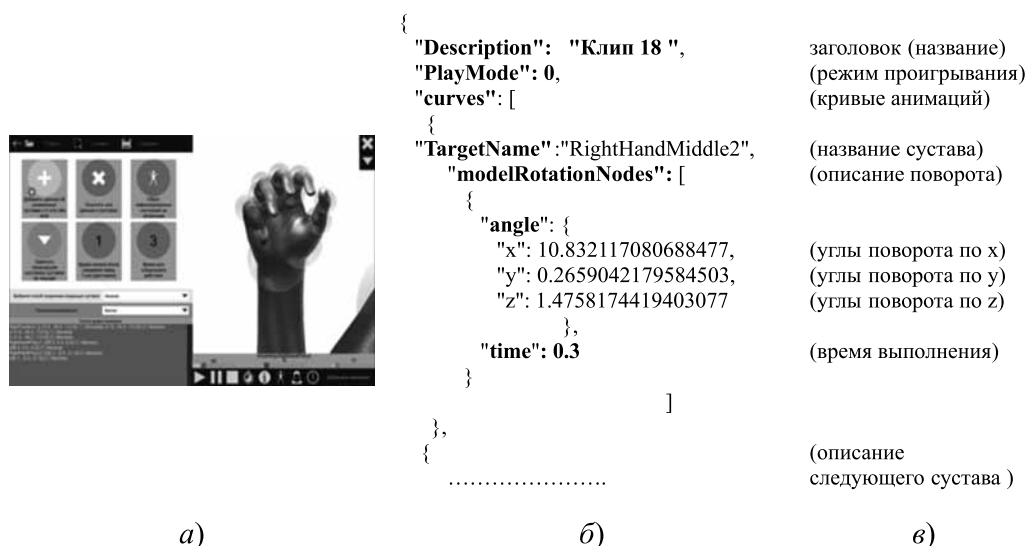


Рис. 2. Редактор клипов (а), анимационный файл на метаязыке (б) и комментарий (в)

### Режим тестирования

Разработанная автором интегрированная система обучения ЖЯ имеет два режима: демонстрация дактилем, слов и высказываний ЖЯ (справочник и уроки); проверка знаний (тестовый режим). Для реализации тестового режима в системе используются следующие модули: редактор тестов; библиотека тестов и эталонов; модуль тестирования. Тестовый режим имеет два варианта проверки

знаний. Первый вариант: аватар демонстрирует жест (слово, дактилему, высказывание), а пользователь должен выбрать из списка правильный ответ (рис. 5). Второй вариант: система просит показать жест, соответствующий дактилеме или слову, и пользователь должен показать его самостоятельно. Второй вариант тестирования опирается на распознавание жеста пользователя.

В редакторе тестов для первого варианта проверки знаний создается библиотека вопросов,

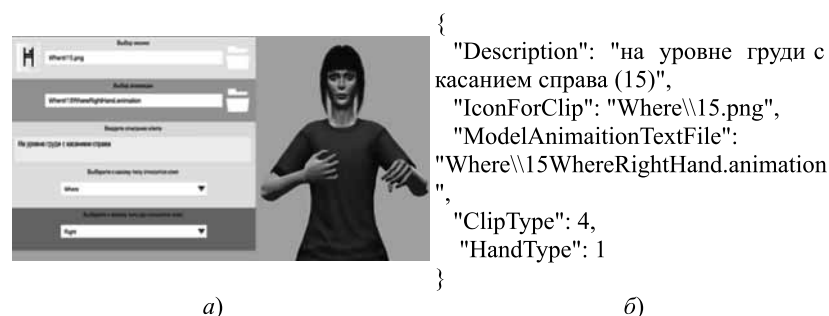


Рис. 3. Редактор знаков нотации (а) и описание символа Димскис на метаязыке (б)

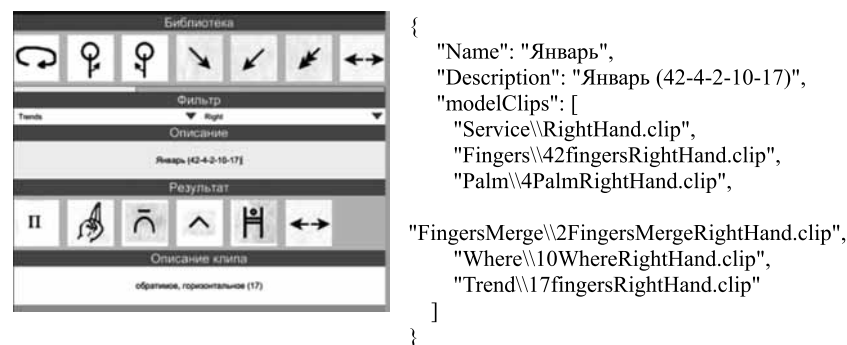


Рис. 4. Пример нотационной записи слова "Январь" и анимационного сценария

которая может содержать дактилемы, слова и высказывания. Редактор доступен через сеть Интернет и сторонним разработчикам.

Для второго из перечисленных выше вариантов тестирования в библиотеку вопросов загружаются эталоны с соответствующими атрибутами, которые получены при демонстрации жестов РЖЯ аватаром.

К разрабатываемой системе специальных требований не предъявляется. Для распознавания жеста используется стандартная веб-камера, поэтому второй вариант тестирования реализуется только на стационарных устройствах с неподвижной веб-камерой.

Распознавание жеста складывается из двух этапов: обнаружение руки на изображении и ее локализация (поиск области интересов); непосредственно распознавание и классификация жеста (сравнение с эталонными образцами по выделенным характерным признакам жеста).



Рис. 5. Режим тестирования с выбором ответа

Требований к окружающей обстановке и специальных требований к освещенности не предусматривается. По этой причине сложно использовать метод выделения жеста по цвету кожи, не усложняя систему. Выполнение тестового вопроса занимает несколько секунд. В течение этого промежутка времени освещенность и окружающую обстановку можно считать неизменными. По этой причине наиболее эффективным методом предварительной обработки изображения является удаление фона. Также в течение этого времени подвижных объектов в кадре, кроме руки, не будет. Таким образом, для обнаружения жеста используется метод детектирования движения.

При переходе в режим тестирования включается веб-камера. В режиме тестирования обучающегося система просит показать определенный жест. С этого момента (или по нажатию клавиши) начинается видеозахват изображения. Система фиксирует первый кадр, который используется для регистрации фона. После этого включается детектор движения. Если соседние кадры отличаются больше, чем на определенное значение  $T$ , то фиксируется начало движения. Детектор переключается в режим фиксации окончания движения. Если соседние кадры не отличаются больше, чем на  $T$ , то обучающийся показывает ключевой жест. Зафиксированный ключевой кадр записывается. Детектор переходит в начальное состояние для определения нового движения, т. е. переключается на фиксацию следующего ключевого кадра. Для статических жестов фиксируется один ключевой кадр, а для динамических — несколько. На демонстрацию жеста отводится несколько секунд или окончание жеста фиксируется нажатием кнопки. По завершению демонстрации жеста записывается

финальный кадр. Он используется для устранения остатков фона. Это связано с тем, что пользователь во время демонстрации может двигаться, и, следовательно, первый и последний кадры будут несколько различаться.

По окончании записи начинается этап локализации жеста. Из всех зафиксированных кадров вычитается первый (фон). Из ключевых кадров вычитается последний кадр, над которым предварительно проведена операция "размытия" в целях надежного устранения остатков фона. В результате остаются ладонь, рука и остатки фона. Наибольшую площадь

с наибольшей яркостью занимает кисть и, возможно, часть руки, что зависит от одежды пользователя.

Для удаления контрастных областей строится гистограмма в градациях серого цвета и используется модифицированный метод Оцу [13]. Для удаления мелких остатков шума применяется метод фильтрации по яркости и связности. Изображение разбивается на  $N$  частей по осям  $X$  и  $Y$ , в каждой части вычисляется число ярких точек. Проверяется связность соседних частей. Такой анализ позволяет исключить разрыв контура изображения руки на этапе локализации. Удаляются области, в которых таких точек меньше  $Z$  ( $Z$  — пороговое значение) и которые являются не связанными или плохо связанными (пороговое значение —  $S$ ). Локализованное изображение области интересов заключается в прямоугольник, и кадр обрезается до его размеров. При этом для динамических жестов фиксируются координаты прямоугольника на исходном кадре. Алгоритм локализации жеста устойчиво работает при слабом освещении как окружающей обстановки (фона), так и рабочего места пользователя. На этом этап локализации жеста заканчивается, дальнейшие действия связаны со следующим этапом — с идентификацией жеста, сравнением локализованного изображения руки обучающегося с подобным изображением аватара.

В ЖЯ основную информацию о жесте несет конфигурация пальцев, поэтому идентификация жеста проводится в два этапа. На первом этапе анализируется кисть целиком. Если на первом этапе принять решение затруднительно, то выполняется второй этап, на котором анализируются только пальцы.

Из локализованного изображения руки выделяется кисть. При разрешении фиксируемого изображения  $640 \times 480$  пикселей кисть занимает не

более 200 пикселей по вертикали или горизонтали в зависимости от демонстрируемого жеста. Это позволяет еще уменьшить размер исследуемой области и при этом удалить малоинформативную часть руки. Изображение выделенной кисти пользователя переводится в градации серого, и строятся две проекции по яркости на осях  $X$  и  $Y$ . При этом проводится дополнительная фильтрация за счет отсеечения не связанных или слабо связанных областей проекций. Такими могут являться часть лица пользователя. Применение метода проекций для сравнения кисти пользователя и аватара позволяет определить качество изображения. Проводится сравнение проекций кисти и эталона, хранящегося в библиотеке тестов. Для этого проекции приводятся к одному масштабу (рис. 6).

Вычисляется корреляция изображений по проекциям. Полученные значения приводятся к интервалу  $[0,1]$  и определяется вероятность совпадения жестов по осям  $X$  и  $Y$  (табл. 1).

Полученные результаты подобия проекций жеста и эталона обрабатываются статистическими методами для принятия решения, которые применяются на нескольких этапах идентификации.

Для принятия решения о правильности показанного жеста используется правило Байеса на основе апостериорной вероятности:

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2), \tag{1}$$

где  $P(A)$  — полная вероятность события  $A$  (совпадение жестов);  $P(B_i)$  — априорная вероятность события  $B_i$  (совпадение по осям  $X$  и  $Y$ );  $P(A|B_i)$  — апостериорная вероятность (результаты сравнения).

Априорные вероятности являются независимыми и имеют нормальное распределение, поэтому равны 0,5:

$$P(A) = 0,5 \times 0,818 + 0,5 \times 0,9757 = 0,8969.$$

При локализованном изображении жеста и завленном эталоне  $S$  система идентификации должна выбрать одну из следующих гипотез:

$H_S$  — жест определяется как эталон  $S$ ;

$H_{\bar{S}}$  — жест не определяется как эталон  $S$ .  
Решение о выборе между двумя гипотезами основано на отношении правдоподобия:

$$W(X) = \frac{p(X | H_S)}{p(X | H_{\bar{S}})} = \frac{P(A)}{1 - P(A)} = \begin{cases} > s, \text{ совпадает с эталоном} \\ \leq s, \text{ не совпадает с эталоном} \end{cases} \tag{2}$$

Величина  $s$  вводится как порог принятия или отклонения гипотезы.

$$W(X) = 0,89689/0,1031 = 8,699.$$

Если имеются сомнения в правильности идентификации и принять окончательное решение затруднительно, то проводится второй этап.

Локализованное изображение жеста обрезается наполовину по высоте при вертикальном жесте и наполовину по ширине при горизонтальном жесте. Также обрезается эталон (извлекается из библиотеки эталонов). Отсчет ведется со стороны пальцев (сверху и справа). По полученным изображениям строятся проекции по вертикали и горизонтали в одинаковом масштабе (рис. 7).

Вычисляется совпадение изображений по проекциям (табл. 2).

Вычисляется вероятность совпадения изображения с эталоном и вычисляется отношение правдоподобия по формулам (1) и (2):

$$P(A) = 0,8190287960, W(X) = 4,5257409928.$$

Таблица 1

Результаты сравнения проекций жеста и эталона для дактилемы  $H$

Проекции	Приведенные значения
Вертикальная по оси $X$	0,8181264923
По оси $Y$ только пальцы	0,9756716953

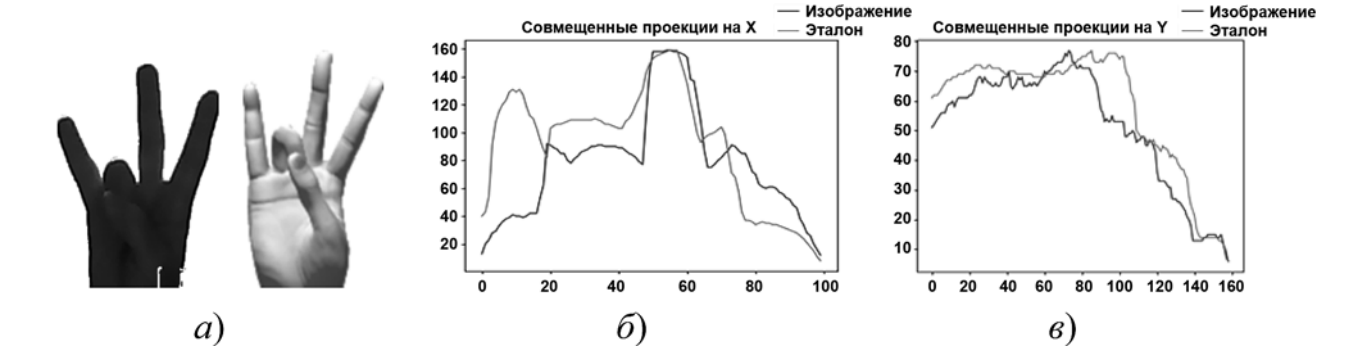


Рис. 6. Изображения кисти пользователя и аватара (а), их совмещенные проекции по оси  $X$  (б) и по оси  $Y$  (в) для дактилемы  $H$

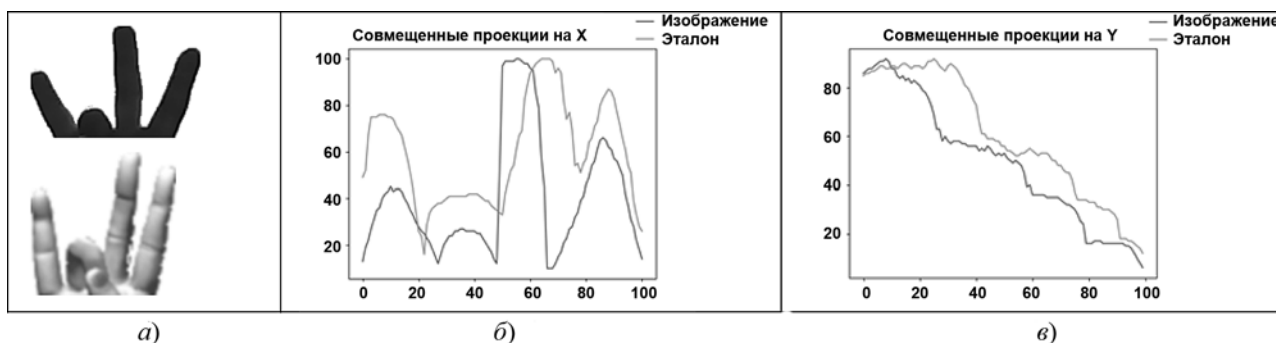


Рис. 7. Изображения пальцев пользователя и аватара, их совмещенные проекции:  
а — изображения; б — вертикальная проекция; в — горизонтальная проекция

На этом этапе принимается окончательное решение с учетом результатов предыдущего этапа.

Рассматривались все возможные варианты локализованного жеста: изображения разного качества, неправильно показанные жесты, вертикальные и горизонтальные жесты, жесты с различным наклоном. Определены границы принятия решения. Значение порога принятия решения зависит от этапов идентификации. В результате проведения экспериментов определены пороговые значения.

Алгоритм принятия решения определяется по следующим действиям и может состоять из двух этапов.

На этапе 1:

1) если значение отношения правдоподобия  $W(X) \leq 3$  ( $s_1 = 3$ ), то жест не распознан по причине плохого качества;

2) если значение отношения правдоподобия  $W(X) > 6$  ( $s_2 = 6$ ), то имеем отличное качество и жест идентифицируется как правильный;

3) если значение отношения правдоподобия  $3 < W(X) \leq 6$  (находится в интервале между значениями 3 и 6), то продолжаем исследование; переход на 2 этап.

На этапе 2:

1) если значение отношения правдоподобия  $W(X) \leq 3$  ( $s_3 = 3$ ), то жест идентифицируется как неверный;

2) если значение отношения правдоподобия  $W(X) > 3$ , то жест идентифицируется как правильный.

Для динамического жеста между ключевыми кадрами определяется тренд перемещения кисти. Сравниваются координаты центров кисти соседних ключевых кадров, вычисляется тренд по горизонтали и вертикали (по координатам  $X$  и  $Y$ ), который сравнивается с эталонным трендом (с ограничениями и неравенствами эталона).

Неравенства формулируются как отношения "больше" или "меньше" в зависимости от движения вниз или вверх:

Таблица 2

Показатели подобия

Проекция	Приведенные значения
Вертикальная по оси $X$	0,6694093655
Горизонтальная по оси $Y$	0,9686482266

$$(X_k - X_n) > e_r \text{ или } (X_k - X_n) < -e_r \\ \text{и } (Y_k - Y_n) > e_b \text{ или } (Y_k - Y_n) < -e_b,$$

где  $X_n$ ,  $Y_n$  — координаты центра кисти первого ключевого кадра;  $X_k$ ,  $Y_k$  — координаты центра кисти второго ключевого кадра;  $e_r$ ,  $e_b$  — допустимые отклонения по горизонтали и вертикали.

В результате проведения экспериментов определены пороговые значения  $e_r = 20$  и  $e_b = 15$  пикселей.

Ограничения используются при повороте кисти, при изменении конфигурации пальцев, при строго горизонтальном или вертикальном движении (кисть остается на месте по соответствующим координатам):

$$|X_k - X_n| < e_r \text{ и/или } |Y_k - Y_n| < e_b.$$

Пример распознавания динамической дактилемы "Щ" показан на рис. 8 и 9.

Вычисляются координаты центра кисти в первом ( $x_1 = 228$ ,  $y_1 = 229$ ) и втором ( $x_2 = 238$ ,  $y_2 = 278$ ) ключевых кадрах.

Вычисляется перемещение:

$$(x_2 - x_1) = |238 - 228| = 10 < e_r, \\ (y_2 - y_1) = (278 - 229) = 49 > e_b.$$

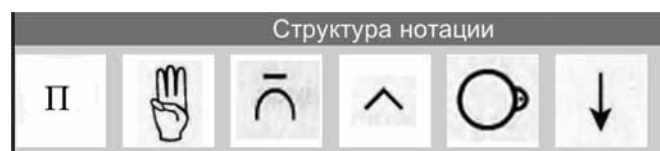


Рис. 8. Структура нотации дактилемы "Щ"

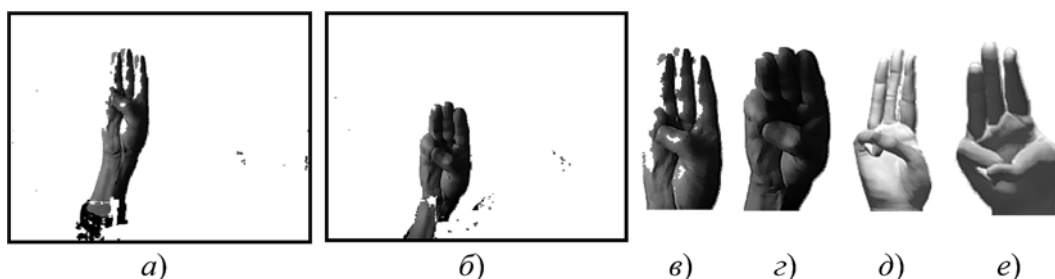


Рис. 9. Первый (а) и второй (б) ключевые кадры после удаления фона на полном растре, локализованные жесты (в) и (з) и соответствующие эталоны (д) и (е)

Определяется тип движения: "перемещение кисти вертикально вниз". Проводится сравнение проекций жестов и эталонов, вычисляются вероятности совпадения изображений с эталонами и определяются отношения правдоподобия ( $W_1(X) = 10,6073739989$ ,  $W_2(X) = 6,5354590156$ ): "оба ключевых кадра определены как совпадающие с эталонами". Итоговый результат: "Жест идентифицируется как правильный".

Для анализа эффективности распознавания жестов в режиме тестирования в интегрированной системе обучения ЖЯ были проведены экспериментальные исследования разработанных средств. Предложенный метод идентификации жеста показывает результат правильного распознавания 95,2 %.

### Заключение

Разработанная автором интегрированная система обучения ЖЯ обладает полным функционалом для создания справочника и уроков ЖЯ. Предложенный метаязык для описания движения подвижных элементов аватара обладает наглядностью, удобством редактирования и позволяет сократить объем файлов более чем на 300 % по сравнению с анимационными файлами Unity 3D.

Система работает в реальном времени и доступна как через сеть Интернет, так и на персональных устройствах. Методика наполнения справочника позволяет сторонним пользователям легко дополнять справочник. Разработаны средства для тестирования знаний пользователя на мобильных устройствах и на стационарных устройствах с неподвижной веб-камерой.

*Автор выражает благодарность научному руководителю работы д-ру техн. наук, проф. кафедры АСУ НГТУ Грифу Михаилу Геннадьевичу.*

### Список литературы

1. O'hara K., Gonzalez G., Sellen A. I. et al. Touchless interaction in surgery // Communications of the ACM. 2014. Vol. 57, No. 1. P. 70—77. DOI: 10.1145/2541883.2541899.
2. Hartmann F., Schlaefer A. Feasibility of touch-less control of operating room lights // International Journal of Computer Assisted Radiology and Surgery. 2012. Vol. 8. P. 259—268. DOI: 10.1007/s11548-012-0778-2.
3. Hongyi Liu, Lihui Wang Gesture recognition for human-robot collaboration // A review International Journal of Industrial Ergonomics Elsevier. 2018. Vol. 68. P. 355—367. DOI:10.1016/J.ERGON.2017.02.004.
4. Hartmann F., Schlaefer A. Feasibility of touch-less control of operating room lights // International Journal of Computer Assisted Radiology and Surgery. 2012. Vol. 8, No. 2. P. 259—268. DOI: 10.1007/s11548-012-0778-2.
5. Кагиров И. А., Рюмин Д. А., Аксенов А. А., Карпов А. А. Мультимедийная база данных жестов русского жестового языка в трехмерном формате // Вопросы языкознания. 2020. № 1. С. 104—123. DOI: 10.31857/S0373658X0008302-1.
6. Карпов А. А., Кагиров И. А. Формализация лексикона системы компьютерного синтеза языка жестов // Труды СПИ-ИРАН. 2011. Вып. 1 (16). С. 123—140. DOI: 10.15622/sp.16.4
7. Гриф М. Г., Лукоянычев А. В. 3D-анимация русского жестового языка на основе нотации Димскис // Программная инженерия. 2017. Т. 8, № 7. С. 310—318. DOI: 10.17587/prin.8.310-318.
8. Мясоедова М. А. Мясоедова З. П. Жестовые нотации и их сравнительный анализ // Современные информационные технологии и ИТ-образование. 2018. [S.l.]. Т. 14. № 1. С. 183—192. DOI: 10.25559/SITITO.14.201801.183-192.
9. Димскис Л. С. Изучаем жестовый язык: учебное пособие. М.: Академия, 2002. 128 с.
10. Речицкая Е. Г. Дактильная и жестовая речь как средства коммуникации лиц с нарушением слуха: учебно-методическое пособие: в 2 частях. М.: Московский педагогический государственный университет (МПГУ), 2016. 144 с.
11. Buttussi F., Chittaro L., Coppo M. Using Web3D technologies for visualization and search of signs in an international sign language dictionary. ACM Press, 2007. 61 p.
12. Гриф М. Г., Лукоянычев А. В. Программный комплекс для обучения русскому жестовому языку на основе Unity3D // Программная инженерия. 2018. Т. 9, № 8. С. 375—384. DOI: 10.17587/prin.9.375-384.
13. Otsu N. A Threshold Selection Method from Gray-Level Histograms // IEEE Transactions on Systems, Man, and Cybernetics. 1979. Vol. 9, No. 1. P. 62—66.



# Integrated Learning System of Sign Language

A. V. Lukoyanychev, dizzystyle@yandex.com, Novosibirsk state technical University, Novosibirsk, 630073, Russian Federation

Corresponding author:

Alexey V. Lukoyanychev, Undergraduate,  
Novosibirsk state technical University, Novosibirsk, 630073, Russian Federation  
E-mail: dizzystyle@yandex.com

Received on June 10, 2022

Accepted on June 27, 2022

The developed integrated system of sign language teaching, aimed at a wide range of users, is considered. The Dimskis notation for sign language writing implementation and the Unity 3D cross-platform development to control an animated 3D character are used in this system. The approach for the describing of 3D character movements based on the meta-language is proposed. The Dimskis notation has been expanded, within the framework of the meta-language. Additional and service characters, control words, special coordinates were introduced. Special control symbols and rules for processes synchronization have been developed. It allows one to implement the movement of several elements at different times, to create complex movements, returning them to the initial state, to cancel the transferences of individual moving elements by using the "fictitious" movement, to change only individual coordinates of the avatar's hand. The technique of creating the sign language reference book is considered. The advantages of this approach are analyzed. In addition to this full-fledged functional toolkit for creating the reference book the integrated system has the ability to test the user's knowledge. The mode gesture demonstration — answer selection is proposed for mobile devices. For stationary devices with a fixed web camera the user has to show the gesture proposed by the system. The following stages of this mode implementation are considered: image capture, gesture localization, gesture identification, decision making. The gesture localization is based on the double background subtraction and histogram method. The gesture identification is based on the projection method and is carried out in two stages. At the first stage the entire hand is analyzed. If it is difficult to make a decision at this stage, then the second stage is carried out, where only fingers are analyzed. The trend of arm movement between key frames is additionally determined for dynamic gestures. The features of each of the stages are shown. The algorithm for making a decision is given. The results of an experimental investigation are presented.

**Keywords:** sign language, Dimskis notation, Unity 3D, avatar control meta-language, gesture localization, gesture identification

For citation:

Lukoyanychev A. V. Integrated Learning System of Sign Language, *Programmnaya Ingeneria*, 2022, vol. 13, no. 8, pp. 404—412.

DOI: 10.17587/prin.13.404-412

## References

1. O'hara K., Gonzalez G., Sellen A. et al. Touchless interaction in surgery, *Communications of the ACM*, 2014, vol. 57, no. 1, pp. 70—77. DOI: 10.1145/2541883.2541899.
2. Hartmann F., Schlaefer A. Feasibility of touch-less control of operating room lights, *International Journal of Computer Assisted Radiology and Surgery*, 2012, vol. 8, pp. 259—268. DOI: 10.1007/s11548-012-0778-2.
3. Hongyi Liu, Lihui Wang Gesture recognition for human-robot collaboration A review *International Journal of Industrial Ergonomics Elsevier*, 2018, vol. 68, pp 355–367. DOI:10.1016/J.ERGON.2017.02.004.
4. Hartmann F., Schlaefer A. Feasibility of touch-less control of operating room lights, *International Journal of Computer Assisted Radiology and Surgery*, 2012, vol. 8, no. 2, pp 259–268. DOI: 10.1007/s11548-012-0778-2.
5. Kagiroy I. A., Ryumin D. A., Aksenov A. A., Karpov A. A. Multimedia database of Russian sign language gestures in three-dimensional format, *Voprosy yazykoznaniya*, 2020, no. 1, pp. 104—123. DOI: 10.31857/S0373658X0008302-1 (in Russian).
6. Karpov A. A., Kagiroy I. A. The formalization of the lexicon of computer synthesis of sign language system, *Trudy SPIIRAN*, 2011, iss. 1 (16), pp. 123—140. DOI: 10.15622/sp.16.4.
7. Grif M. G., Lukoyanychev A.V. 3D - animation of Russian sign language on the basis of Dimskis notation, *Programmnaya Ingeneria*, 2017, vol. 8, no. 7, pp. 310–318. DOI: 10.17587/prin.8.310-318. (in Russian).
8. Myasoedova M. A., Myasoedova Z. P. Sign notations and their comparative analysis, *Sovremennye informacionnye tekhnologii i IT-obrazovanie*, 2018, [S.l.], vol. 14, no. 1, pp. 183—192. DOI: 10.25559/SITITO.14.201801.183-192 (in Russian).
9. Dimskis L. S. *The study of sign language: study guide*, Moscow, Academy, 2002, 128 p. (in Russian).
10. Rechitskaya E. G. *Dactyl and sign language as a means of communication of persons with hearing impairment: educational and methodical manual*: in 2 parts, Moscow, Moscow Pedagogical State University (MPSU), 2016, 144 p. (in Russian).
11. Buttussi F., Chittaro L., Coppo M. *Using Web3D technologies for visualization and search of signs in an international sign language dictionary*, ACM Press, 2007, 61 p.
12. Grif M. G., Lukoyanychev A. V. The program complex for training russian sign language based on Unity3D, *Programmnaya Ingeneria*, 2018, vol. 9, no. 8, pp. 375—384. DOI: 10.17587/prin.9.375-384 (in Russian).
13. Otsu N. A Threshold Selection Method from Gray-Level Histograms, *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, vol. 9, no. 1, pp. 62—66.

ООО "Издательство "Новые технологии". 107076, Москва, ул. Матросская Тишина, д. 23, стр. 2  
Технический редактор Е. М. Патрушева. Корректор А. В. Чугунова.

Сдано в набор 21.07.2022 г. Подписано в печать 24.08.2022 г. Формат 60×88 1/8. Заказ PI821  
Цена свободная.

Оригинал-макет ООО "Авансд солишнз". Отпечатано в ООО "Авансд солишнз".  
119071, г. Москва, Ленинский пр-т, д. 19, стр. 1. Сайт: [www.aov.ru](http://www.aov.ru)