

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

G09B 21/00 (2024.08)

(21)(22) Заявка: 2024111207, 23.04.2024

(24) Дата начала отсчета срока действия патента:
23.04.2024Дата регистрации:
11.03.2025

Приоритет(ы):

(22) Дата подачи заявки: 23.04.2024

(45) Опубликовано: 11.03.2025 Бюл. № 8

Адрес для переписки:

124460, Москва г. Зеленоград, а/я 200, ООО
"Институт инноваций и права"

(72) Автор(ы):

Харламенков Алексей Евгеньевич (RU)

(73) Патентообладатель(и):

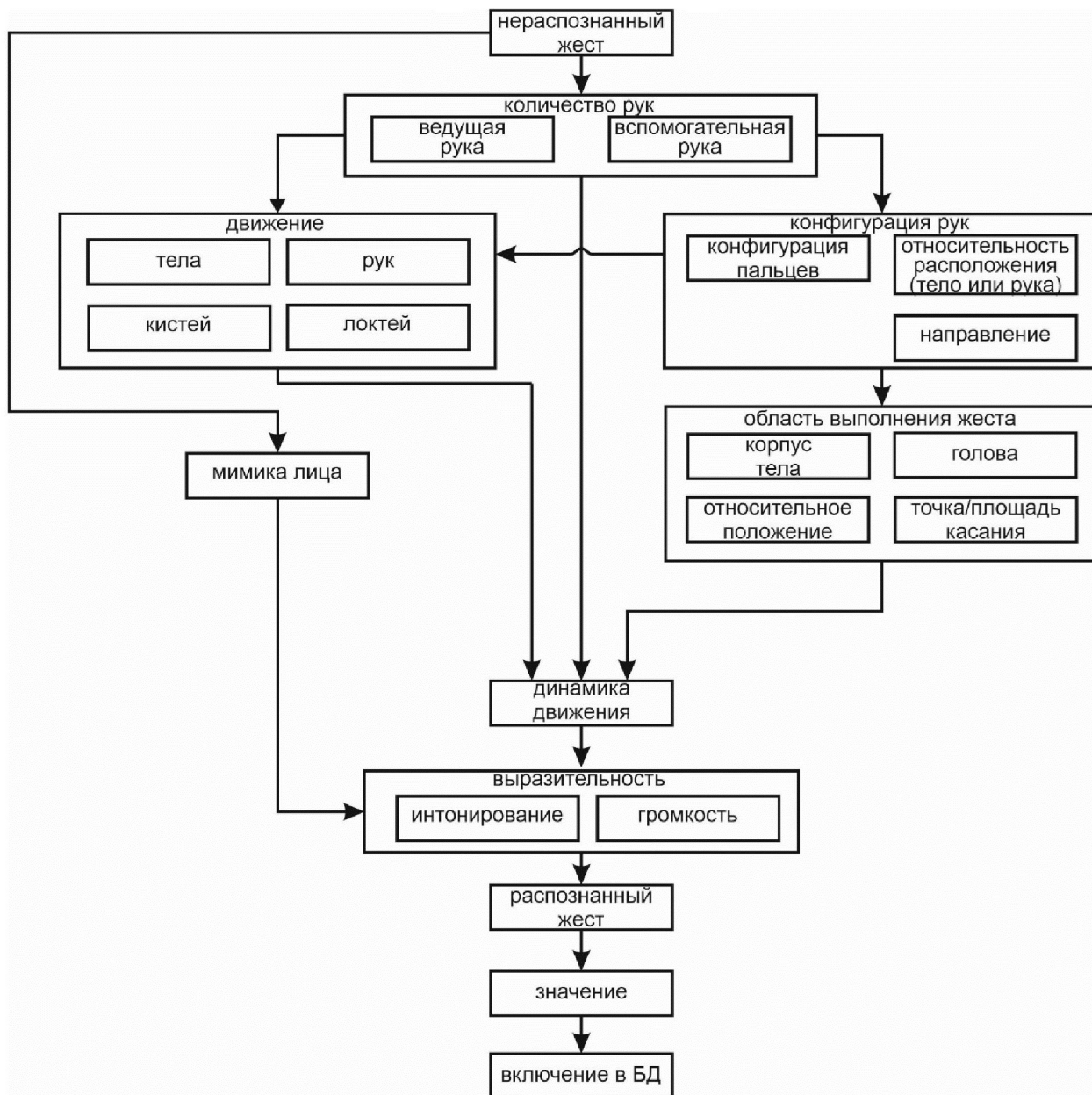
Харламенков Алексей Евгеньевич (RU)

(56) Список документов, цитированных в отчете
о поиске: US 6116907 A, 12.09.2000. US
2022414350 A1, 29.12.2022. US 2020286493 A1,
10.09.2020. US 10129395 B1, 13.11.2018.(54) СПОСОБ ГЕНЕРАЦИИ И РАСПОЗНАВАНИЯ ЖЕСТОВ ЖЕСТОВОГО ЯЗЫКА И УСТРОЙСТВО
ДЛЯ ЕГО РЕАЛИЗАЦИИ

(57) Реферат:

Изобретение относится к способам и системам распознавания жестов. Технический результат - повышение точности распознавания жеста жестового языка, в том числе с учетом индивидуального почерка жестикулирующего, - достигается за счет того, что формируют массив данных содержащий матрицу жестов и связанную с матрицей жестов матрицу кирем жестов, включающую киремы и их конфигурации, для чего осуществляют захват, разделение жеста на киремы, из которых формируют матрицу кирем, каждой из кирем присваивают уникальное обозначение для последующей идентификации киремы в матрице кирем и распознавания жеста в связанной с матрицей кирем матрице жестов, при этом при разделении жеста на киремы учитывают их последовательность, количество, распределение во времени и во взаимном временном расположении относительно друг друга в жесте и представляют в виде реляционной формулы для последующей генерации жеста путем подачи формулы жеста на вход

графической модели человека и воспроизведения этой графической моделью человека жеста в жестовой вселенной. Для распознавания жеста захватывают изображение жеста, осуществляют разделение жеста на киремы, которые сравнивают с киремами в матрице кирем и осуществляют идентификацию кирем, при этом при отсутствии в матрице кирем уточняют значение этой киремы и если кирема является новой, то ее включают в матрицу кирем в качестве эталонной после кирематического разбора и присвоения ей уникального обозначения, если показываемая кирема является вариантом одной из существующих в матрице кирем, то ее включают в матрицу кирем в качестве варианта существующей киремы с присвоением ей уникального обозначения существующей киремы, если показываемая кирема по результатам кирематического разбора жеста признана ошибочной в показе, корректируют эту кирему в жесте и повторно воспроизводят и захватывают для распознавания. 2 н. и 5 з.п. ф-лы, 8 ил.



Фиг.1



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(12) **ABSTRACT OF INVENTION**

(52) CPC
G09B 21/00 (2024.08)

(21)(22) Application: **2024111207, 23.04.2024**

(24) Effective date for property rights:
23.04.2024

Registration date:
11.03.2025

Priority:

(22) Date of filing: **23.04.2024**

(45) Date of publication: **11.03.2025** Bull. № 8

Mail address:

**124460, Moskva g. Zelenograd, a/ya 200, OOO
"Institut innovatsij i prava"**

(72) Inventor(s):

Kharlamenkov Aleksei Evgenevich (RU)

(73) Proprietor(s):

Kharlamenkov Aleksei Evgenevich (RU)

(54) **METHOD OF GENERATING AND RECOGNIZING GESTURES OF SIGN LANGUAGE AND DEVICE FOR ITS IMPLEMENTATION**

(57) Abstract:

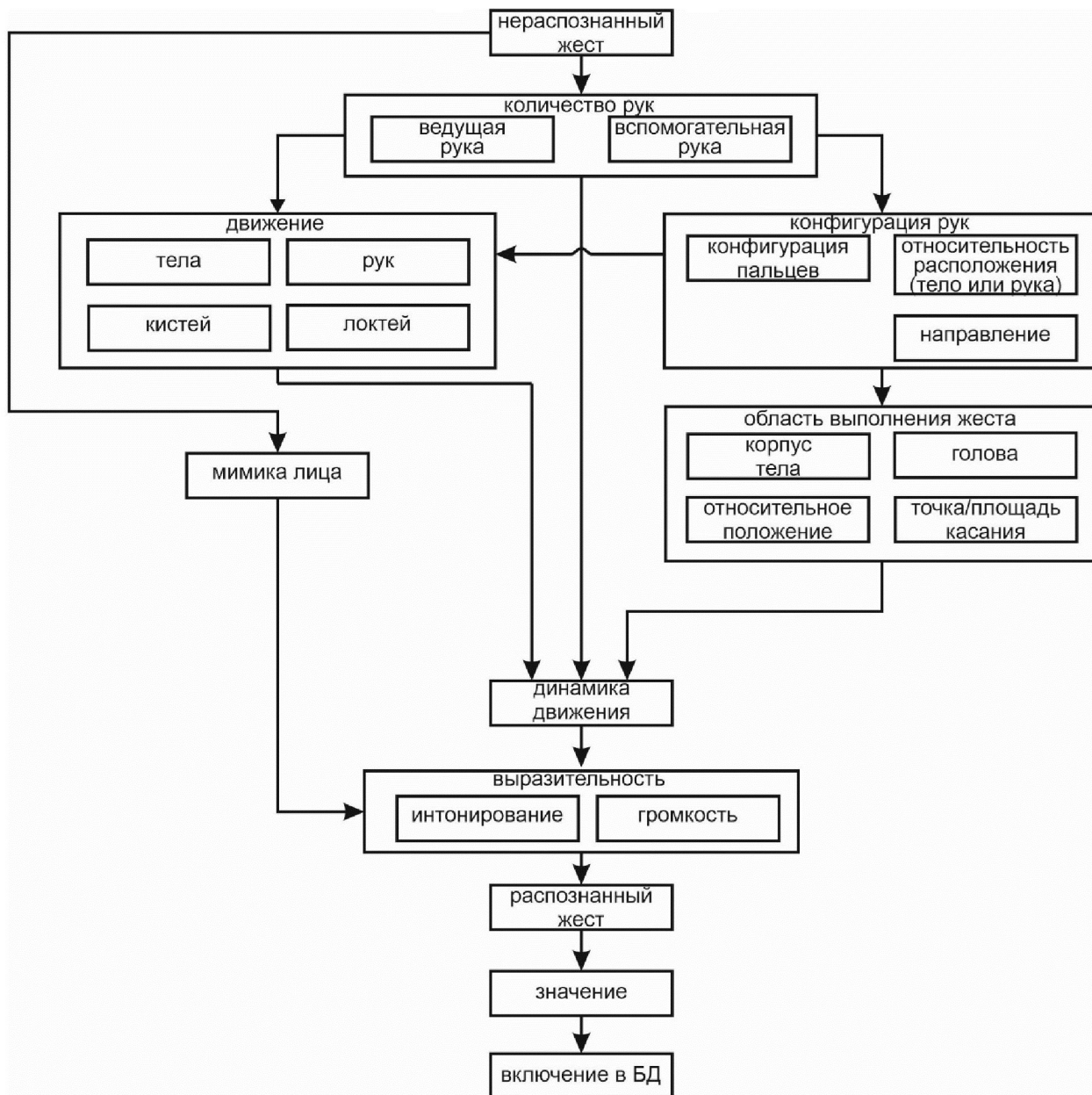
FIELD: data processing.

SUBSTANCE: invention relates to gesture recognition methods and systems. Forming a data array containing a gesture matrix and a gesture matrix associated with the gesture kireme matrix, including the kiremes and their configurations, for this purpose, gesture is gripped, divided into kiremes, from which kireme matrix is formed, each of the kireme is assigned a unique designation for the subsequent identification of the kireme in the kireme matrix and the gesture recognition in the gesture matrix associated with the kireme matrix, when dividing a gesture into kiremes, their sequence, number, distribution in time and in mutual temporal arrangement relative to each other in a gesture are taken into account and presented in the form of a relational formula for subsequent gesture generation by sending a gesture formula to the input of a graphic model of a person and reproducing the gesture in the gesture world by this graphic model. To recognize a gesture, a gesture image is captured, the gesture is

divided into kiremes, which are compared with the kiremes in the kireme matrix and the kiremes are identified, at that, in the absence in the kireme matrix, the meaning of this kireme is specified and if the kireme is new, then it is included in the kireme matrix as a reference after the kirematic analysis and assignment of a unique designation to it, if the displayed kireme is a version of one of the kiremes existing in the matrix, then it is included in the kireme matrix as a version of the existing kireme with the assignment of a unique designation of the existing kireme, if the displayed kireme is recognized as an error in the display based on the results of the kirematic analysis of the gesture, this kireme is corrected in the gesture and repeated and captured for recognition.

EFFECT: high accuracy of recognizing a gesture of sign language, including taking into account the individual handwriting of the sign language.

7 cl, 8 dwg



Фиг.1

Изобретение относится к способам и системам распознавания жестов [G09B 21/00].

Из уровня техники известно УСТРОЙСТВО ДЛЯ ПЕРЕВОДА ЯЗЫКА ЖЕСТОВ [JPН08328460 (А), опубликовано: 13.12.1996], предназначенное для хранения

стандартного изображения руки, которое представляет собой изображение руки, включающее точки, указывающие положения обоих плеч и локтей при воспроизведении элементов языка жестов, а также положения и формы обеих рук, вместе с идентификатором для каждого элемента жестового языка, таблицу преобразования, хранящую словесную информацию, соответствующую идентификатору стандартного изображения руки, и словесную информацию, соответствующую комбинации идентификаторов множества стандартных изображений руки, вход для ввода неподвижного изображения знака элемента языка в состоянии языка жестов жестикулирующего лица, при этом изображение, соответствующее конкретному элементу языка жестов, сохраненному в устройстве, получают заранее, и изображение руки, подаваемое из устройства, стандартизируется посредством значения коррекции стандартизации средством стандартизации для создания идентификатора путем последовательного поиска идентификатора стандартного изображения руки, которое аппроксимирует стандартизированное изображение руки, подаваемое из средства стандартизации, и ответа на команду преобразования, подаваемую извне, и средство распознавания для поиска в таблице преобразования для информации о слове, соответствующей этой строке идентификатора, и средство вывода для вывода соответствующего голоса и отображения соответствующего дисплея на основе информации о слове.

Недостатком аналога, является низкая точность распознавания жеста, обусловленное тем, что в заявленной системе каждый из жестов представляет собой неделимый объект, который распознается в целом по совокупности точек скелета и не предполагает собой разделение на составные элементы, присущие каждому жесту. Следствием данного недостатка является невозможность нивелирования жестового почерка при распознавании жестов.

Наиболее близкой по технической сущности является СИСТЕМА КОДИРОВАНИЯ И ПОЛУЧЕНИЯ ЗРИТЕЛЬНЫХ СИГНАЛОВ [US6116907 (А), опубликовано:

12.09.2000], где система, кодирующая знаки на заданном языке жестов, использующая по меньшей мере первую руку, причем указанная система содержит средство для определения конфигурации первой руки, средство для определения ориентации ладони первой руки, средство для определения местоположения первой руки, средство для определения движения первой руки, средство кодирования, соединенное с каждым из средств для определения конфигурации первой руки, ориентации ладони, местоположения и движения соответствующего кодового значения, средство объединения, соединенное со средством кодирования, для объединения в комбинированное значений кодовых значений конфигурации первой руки, ориентации ладони, местоположения и движения и средство выбора, соединенное со средством объединения, для выбора одного или более соответствующих знаков с использованием значений комбинированного кода, причем указанное средство выбора работает для сопоставления комбинированному коду значения по базе данных кодовых значений и возвращать из указанной базы данных кодовых значений количество объединенных кодовых значений одного или нескольких знаков, попадающих в заданный диапазон кодовых значений.

Основной технической проблемой аналога и прототипа является малое количество кистей, на которые разбит жест жестового языка, что ограничивает возможность

детектирования жеста полностью, что приводит к ошибкам нахождения распознаваемого жеста в базе данных и определения его значения. Кроме того, при распознавании жестов в заявленных решениях не учитывается длительности элементов

Задача изобретения состоит в устранении недостатков аналога и прототипа.

Технический результат изобретения заключается в повышении точности распознавания жеста жестового языка, в том числе с учетом индивидуального почерка жестикулирующего.

Указанный технический результат достигается за счет того, что способ генерации и распознавания жестов жестового языка, характеризующийся тем, что первоначально формируют массив данных содержащий матрицу жестов, включающую в себя идентификатор, текстовое наименование, описание жеста, изображение жеста и связанную с матрицей жестов матрицу кирем жестов, включающую киремы и их конфигурации, для чего при воспроизведении жеста осуществляют его захват, разделение на киремы из которых формируют матрицу кирем, каждой из кирем присваивают уникальное обозначение в матрице по группе к которой относится кирема и списку возможных конфигураций этой киремы в этой группе для последующей идентификации киремы в матрице кирем и распознавания жеста в связанной с матрицей кирем матрице жестов, при этом при разделении жеста на киремы учитывают их последовательность, количество, распределение во времени и во взаимном временном расположении относительно друг другу в жесте и представляют в виде реляционной формулы, уникальной для каждого жеста и содержащей последовательность, количество, таймлайн кирем и списочное значение каждой из кирем в ее группе для последующей генерации жеста путем подачи формулы жеста на вход графической модели человека и воспроизведения этой графической моделью человека жеста в принятой для жестовой вселенной системе координат, для распознавания жеста захватывают изображение жеста, осуществляют разделение жеста на киремы, которые сравнивают с киремами в матрице кирем и осуществляют идентификацию кирем, при этом при отсутствии в матрице кирем уточняют значение этой киремы и если кирема является новой, то ее включают в матрицу кирем в качестве эталонной после кирематического разбора и присвоения ей уникального обозначения, если показываемая кирема является вариантом одной из существующих в матрице кирем, то ее включают в матрицу кирем в качестве варианта существующей киремы с присвоением ей уникального обозначения существующей киремы, если показываемая кирема по результатам кирематического разбора жеста признана ошибочной в показе, корректируют эту кирему в жесте и повторно воспроизводят и захватывают для распознавания.

В частности, жест разбивают на киремы, характеризующие количество рук, используемых в жесте, конфигурацию ведущей руки, конфигурацию вспомогательной руки, положение кистей относительно тела, положение кистей относительно друг друга, движение кистей, движение локтей, движение тела, движение рук, область выполнения жеста, мимику лица, интенсивность и амплитуду выполнения жеста, перемещаемость жеста.

В частности, конфигурации вспомогательной руки являются зеркальным отображением конфигураций ведущей руки.

В частности, количество кирем в жесте в процессе изучения или развития жестового языка может увеличиваться.

В частности, для включения жеста в массив данных жестов в качестве эталонного жест воспроизводит профессиональный диктор или пользователь, профессионально владеющий жестовым языком.

Устройство для генерации и распознавания жестов, включающее модуль захвата жеста, к модулю захвата жеста последовательно подключены модуль обнаружения объектов на захватываемом с помощью модуля захвата жеста изображении, модуль отслеживания рук и модуль обнаружения лица, модули обнаружения объектов, 5 отслеживания рук, обнаружения лица подключены к модулю обработки, при этом к каждому из модулей обнаружения объектов, отслеживания рук, обнаружения лица подключены счетчики, выполненные с возможностью фиксации последовательности и длительности каждого из элементов жеста, захваченного с помощью модуля захвата жеста, выход модуля обработки подключен к модулю сравнения, к одному выходу 10 модуля сравнения подключен блок матриц кирем, к другому из выходов модуля сравнения последовательно подключены модуль накопления, модуль идентификации, блок матрицы жестов, интерфейсный модуль и блок ввода данных, к модулю идентификации параллельно подключен генератор чисел, к блоку матриц кирем параллельно подключен блок матриц жестов, интерфейсный модуль подключен к 15 модулю захвата жеста.

В частности, модуль захвата жеста выполнен в виде устройств машинного зрения, например, стереокамер или систем камер и датчиков глубин выполненных с возможностью получать X, Y, Z-координаты жеста.

Краткое описание чертежей.

На фиг. 1 показана блок-схема разложения жеста на киремы.

На фиг. 2 показан вид сверху жестовой вселенной.

На фиг. 3 показан вид спереди жестовой вселенной.

На фиг. 4 показан вид сбоку жестовой вселенной

На фиг. 5 показана комплексная временная шкала выполнения жеста (таймлайн)

25 для каждой из кирем на примере сложносоставного жеста «Абажур».

На фиг. 6 показано устройство для генерации и распознавания жестов.

На фиг. 7 показан алгоритм распознавания жестов.

На фиг. 8 показаны примеры жестов, где слева показан жест, обозначающий букву «В», справа – жест, обозначающий цифру «5».

30 На фигурах обозначено: 1 – модуль захвата изображения, 2 – модуль обнаружения объектов, 3 – модуль отслеживания рук, 4 – модуль обнаружения лица, 5 – модуль обработки, 6 – счетчики, 7 – модуль сравнения, 8 – блок матриц кирем, 9 – модуль накопления, 10 – модуль идентификации, 11 – блок матрицы жестов, 12 – генератор чисел, 13 – интерфейсный модуль, 14 – блок ввода данных.

35 Осуществление изобретения.

В современном мире количество областей, где человек применяет жесты в своей деятельности, увеличивается. Это и управление роботами на производстве, в строительстве, медицине, управление функционалом в автомобиле, телефоне, использование в интерактивных развлечениях и управлении Интернет вещами (IoT).

40 Самым востребованным является жестовый язык для коммуникации между людьми с ограниченными возможностями по слуху. Жест в жестовом языке может быть простым и сложносоставным, но состоящим максимум из трех простых жестов. Если больше, то по теории жестового языка жест, состоящий более чем из трех простых жестов, является жестовым рассказом. При этом жестовый язык и его машинное распознавание 45 является самым сложным в реализации.

Сущность изобретения заключается в разделении жеста на составляющие – «киремы», формировании массива данных из эталонных кирем, составляющих эталонные жесты, сравнении кирем вновь воспроизведенного жеста с эталонными и распознавании вновь

воспроизведенного жеста. Кирема – это минимальная пространственно-кинетическая единица строя жестового языка. Разделение жеста на киремы или нотирование жеста представляет собой процесс создания уникальной формулы жеста.

Русский жестовый язык принадлежит к семейству французских жестовых языков.

5 Лексика русского жестового языка была сформирована австрийским жестовым языком. Русский жестовый язык до сих пор недостаточно описан и относится к языкам с низким уровнем ресурсов. Исследования жестового языка требуют создания базы данных, необходимой для обучения, тестирования и сравнения различных систем распознавания, основанных на глубоком обучении. В любом жестовом языке все жесты состоят из
10 кирем, которые в своей массе идентичны в разных национальных жестовых языках, поэтому предлагаемый способ применим для любого из большинства жестовых языков.

Для русского жестового языка детальный состав кирем проработан и зафиксирован в электронной справочно-аналитической системе (ЭСАС) «Толкового лексикографического словаря русского жестового языка» (<https://slovar.surdocentr.ru>).

15 При машинном распознавании жестового языка, в том числе отличного от русского языка, состав кирем может дополняться в горячем режиме при распознавании жестов соответствующих жестовых языков по предлагаемому в изобретении алгоритму. Поэтому массив данных или база данных русского жестового языка может быть взят за основу распознавания любого жестового языка.

20 В настоящем изобретении жест жестового языка предлагается разбивать на киремы, каждой из которых присваивается в определенной закономерности уникальное обозначение (имя или номер), идентифицирующее кирему в массиве (базе) данных известных кирем и жестов, и состоящий, например, из последовательного набора букв и цифр, где буквы обозначают группу к которой относится кирема, а цифра - это ее
25 номер (численное значение) в этой группе кирем.

Перечень наименований (групп) кирем, на которые разбивают жест жестового языка (см. Фиг. 1):

- 1) количество рук Q_H ;
- 2) конфигурация ведущей руки WL_X ;
- 30 3) конфигурация вспомогательной руки WA_Y является зеркальным отображением конфигураций ведущей руки WL_X ;
- 4) положение (направленность) кистей относительно тела WB_Z ;
- 5) положение кистей относительно друг друга WW_C ;
- 35 6) движение кистей WM_R (круговое, прямолинейное, волновое и т.д.);
- 7) движение локтей EM_U ;
- 8) движение тела (наклоны вперед, назад, влево, вправо, повороты влево, вправо и т.д.) BL ;
- 40 9) движение рук (прямолинейное, по спирали, по дуге, волнообразное и т.д.) NM_J ;
- 10) область выполнения жеста (над головой, перед лицом, перед грудью и т.д.) T_I ;
- 11) мимика лица F – аналог интонации (вопросительная, утвердительная и т.д.);
- 12) интенсивность выполнения жеста I_K (аналог скорости речи);
- 45 13) амплитуда выполнения жеста A_L (аналог громкости речи);
- 14) перемещаемый (неперемещаемый) жест M_P .

Каждая из кирем имеет свой список значений, которое она может принимать в том или ином жесте. По «количеству рук» перечень значений списка первой киремы

следующий: одноручный жест Q_1 ; двуручный жест Q_2 . Список данной киремы исчерпывающий, то есть $N \in [1; 2]$. Аналогично касается киремы M , характеризующей перемещаемый (неперемещаемый) жест, которая имеет в списке P всего два значения – перемещаемый и неперемещаемый.

Для всех остальных кирем списки значений являются открытыми, так как количество выполняемых движений постоянно пополняются.

По конфигурациям киремы, характеризующей конфигурацию ведущей руки WL_X численное значение киремы $X \in [1; n]$, где n – количество известных конфигураций ведущей руки в жестовом языке, учитывающей количество участвующих в киреме WL пальцев, их взаимную конфигурацию, направление их движения, в том числе относительно друг друга или кисти, наличие или отсутствие контакта пальцев (кисти) с лицом или другими частями тела, площадь контакта (пальцем, несколькими пальцами, ладонью кисти – внутренней или наружной стороной и т.д.). В настоящее время известно более 80 конфигураций киремы, характеризующей конфигурацию ведущей руки WL , то есть $X \in [1; 80]$, но оно может увеличиваться в процессе записи новых жестов и развития жестового языка.

Аналогично, для киремы, характеризующей конфигурацию вспомогательной руки WA_Y численное значение киремы $Y \in [1; m]$, где m – количество известных конфигураций вспомогательной руки в жестовом языке, учитывающей количество участвующих в киреме WA пальцев, их взаимную конфигурацию, направление их движения, в том числе относительно друг друга или кисти, наличие или отсутствие контакта пальцев (кисти) с лицом или другими частями тела, площадь контакта (пальцем, несколькими пальцами, ладонью – внутренней или наружной стороной и т.д.).

Список Z киремы, характеризующей положение кистей относительно тела WB определяется расположением кисти (перпендикулярным, параллельным и т.д.) относительно тела.

Список C киремы, характеризующей положение кистей относительно друг друга WW определяется расположением кистей относительно друг друга.

Положение жеста относительно тела или другой кисти (киремы WB или WW) может быть, например, перпендикулярным или параллельным и т.д. Списки Z или C кирем WB или WW может увеличиваться в других жестовых языках или с развитием жестового языка.

Список R киремы, характеризующей движение кистей WM определяется, соответственно, известными вариантами движения кистей в существующих жестах жестового языка, при этом направление движения предусматривает все стороны пространства относительно говорящего, а также диагональные плоскости, что, например, в известных жестах русского жестового языка может составлять в настоящий момент до восьми направлений; с развитием языка может увеличиваться.

Список U киремы, характеризующей движение локтей EM определяется, соответственно, известными вариантами движения локтей в существующих жестах жестового языка, которые определяются либо требованиями обратной кинематики в тех случаях, когда на локте не акцентируется внимание, либо данное движение является акцентным.

Кирема T_I характеризует область выполнения жеста или локацию жеста, центром которой всегда является исполнитель жеста. К примеру, жесты, обозначающие мыслительные процессы выполняют в районе лба; жесты, обозначающие эмоции – в районе груди и т.д. Для генерации и распознавания жеста определяют квадранты

(внутреннее деление объема) и границы жестовой вселенной – пространства, доступного рукам исполнителя жеста, где исполняется жест (см. Фиг.2-4). В настоящем техническом решении под жестовой вселенной понимают пространство вокруг человека, границы которого определены длиной рук. Для определения области выполнения жеста (локации выполнения жеста) пространство для каждой из рук делят на значимые квадранты, такие как, над головой слева, над головой справа, над головой по центру, перед лицом, слева от лица, справа от лица и т.д.). Перемещение жеста из квадранта в квадрант приводит к изменению смысла, а также к образованию другого жеста. Количество квадрантов и границы зависят от разнообразия жестов в языке. В настоящее время в русском жестовом языке известна 21 позиция локации жеста, поэтому и значения из списка этой киремы $I \in [1; 21]$.

Список L киремы, характеризующей движение тела V_L определяется, соответственно, известными вариантами движений тела в существующих жестах жестового языка, например, наклоны вперед, назад, влево, вправо и т.д.

Список J киремы, характеризующей движение рук NM определяется соответственно, известными вариантами движений рук в существующих жестах жестового языка, например, прямолинейное, по спирали, по дуге, волнообразное и т.д.

Мимика лица F выполняет роль интонирования (вопрос, восклицание и т.д.) и носит в жесте необязательный характер.

Интенсивность выполнения жеста I (вялость, резкость, плавность жеста и т.д.) зависит от характера передаваемой информации с помощью жестового языка, ее значимости. Интенсивность жеста также зависит от темперамента исполнителя и может определять особенности индивидуального стиля исполнителя.

Амплитуда выполнения жеста A (компактная, размашистая и т.д.) является прямым аналогом громкости в звуковом языке.

При разделении жеста на киремы учитывают их последовательность и время выполнения, выраженное, как правило, в долях относительно времени выполнения всего жеста.

На фиг.5 показана комплексная временная шкала выполнения жеста (таймлайн) для каждой из кирем на примере сложносоставного жеста «Абазур». Таймлайн – это графическое распределение кирем жеста во времени и во взаимном временном расположении их относительно друг друга, а также последовательность в пределах временной шкалы одной киремы. На основе таймлайна формируется формула жеста. Только таймлайн позволяет ввести в жесте такую характеристику, как время.

Количество кирем, а следовательно и временных дорожек на таймлайне в процессе изучения (развития) жестового языка может увеличиваться.

Последовательность, количество, таймлайн и значение каждой из кирем из списка в жесте, идентифицируемые с помощью уникальных обозначений предлагается представлять в виде реляционной формулы жеста, то есть набора аргументов и значений этих аргументов. Формула, содержащая набор аргументов и значения этих аргументов уникальна для каждого жеста, и может выглядеть в обобщенном виде при последовательном выполнении кирем, указанных в списке, как:

Например, при одновременном выполнении кирем ведущей и вспомогательной руками, обобщенная формула жеста может выглядеть как:

В качестве примера рассмотрим формирование формул простых жестов.

Пример 1. Жест, обозначающий понятие «прямо»

Согласно ЭСАС, жест выполняется одной ведущей правой рукой. Правая кисть в "В-конфигурации" обращена ладонью влево, кончиками пальцев вверх, по прямой

осуществляется движение вперёд этой рукой.

Киремы жеста, обозначающего «прямо», имеют следующие значения (аргументы приведены примерными):

- количество рук Q – используется одна рука, то есть $H = 1$;
- 5 - конфигурация ведущей руки WL – кисть прямая, пальцы прижаты друг к другу, то есть $X = 3$;
- конфигурация вспомогательной руки WA – не участвует, то есть $Y = 0$;
- положение кистей (жеста) относительно тела WB – перпендикулярное, то есть $Z = 1$;
- 10 - положение кистей относительно друг друга (жеста относительно вспомогательной кисти) WW – вторая кисть не участвует, $C = 0$;
- движение кистей WM – вперед от себя, $R = 1$;
- движение локтей EM – зависимое, $A = 1$;
- движение тела B – отсутствует, $L = 0$;
- 15 - движение рук NM – прямолинейное, $J = 1$;
- область выполнения жеста T – на уровне груди, $I = 9$;
- мимика F – отсутствует;
- интенсивность выполнения жеста E – плавная, $K = 3$;
- амплитуда выполнения жеста A – компактная, $L = 1$;
- 20 - жест непереключаемый, то есть для киремы M ее значение $P = 2$.

Формула жеста, обозначающего понятие «прямо» принимает вид:

«прямо» = $Q_1 WL_3 WB_1 EM_1 NM_1 T_9 K_3 L_1 P_2$

Пример 2. Жест, обозначающий понятие «коридор»

Согласно ЭСАС, жест выполняется одновременно ведущей правой и вспомогательной 25 левой руками. Правая и левая кисти в "В-конфигурации" обращены ладонями навстречу друг другу и расположены напротив, кончиками пальцев кверху, по прямой осуществляется одновременное движение вперёд этими руками.

Киремы жеста, обозначающего «коридор», имеют следующие параметры (значения аргументов приведены примерными):

- 30 - количество рук Q – используется две руки, то есть $H = 2$;
- конфигурация ведущей руки WL – кисть прямая, пальцы прижаты друг к другу, то есть $X = 3$;
- конфигурация вспомогательной руки WA – кисть прямая, пальцы прижаты друг к 35 другу, то есть $Y = 3$;
- положение кистей (жеста) относительно тела WB – перпендикулярное, то есть $Z = 1$;
- положение кистей относительно друг друга (жеста относительно вспомогательной кисти) WW – напротив друг друга, $C = 1$;
- 40 - движение кистей WM – вперед от себя, $R = 1$;
- движение локтей EM – зависимое, $A = 1$;
- движение тела B – отсутствует, $L = 0$;
- движение рук NM – прямолинейное, $J = 1$;
- область выполнения жеста T – на уровне груди, $I = 9$;
- 45 - мимика F – отсутствует;
- интенсивность выполнения жеста E – резкая, $K = 2$;
- амплитуда выполнения жеста A – размашистая, $L = 2$;
- жест непереключаемый, то есть для киремы M ее значение $P = 2$.

Формула жеста, обозначающего понятие «прямо» принимает вид:

$$\text{коридор} = Q_2 \frac{WL_3}{WA_3} WB_1 WW_1 WM_1 EM_1 HM_1 T, K_2 L_2 M_2 .$$

Для распознавания жестов первоначально необходимо сформировать массив данных жестов, содержащий матрицу жестов, включающей в себя идентификатор известного жеста, текстовое наименование жеста, описание жеста и его кирем, изображение с реперными точками жеста, видео жеста и матрицы кирем, включающие киремы и их конфигурации (аргументы и значения этих аргументов). Массив данных жестов формируют с помощью устройства для распознавания жестов.

Устройство для генерации и распознавания жестов содержит модуль захвата жеста 1 (см. Фиг. 6), выполненный в виде устройств машинного зрения, например, стереокамер или систем камер и датчиков глубин. Для формирования массива данных эталонных жестов модуль захвата жеста 1 смонтирован в специальном помещении. Например, в качестве помещения можно использовать кубическое помещение со смонтированными в нем двумя веб-камерами с известными положениями и углами, работающими и записывающими одновременно, позволяющими получать X, Y, Z-координаты жеста.

К модулю захвата жеста 1 последовательно подключены модуль обнаружения объектов 2 на захватываемом с помощью модуля захвата жеста 1 изображении, модуль отслеживания рук 3 и модуль обнаружения лица 4.

Модули обнаружения объектов 2, отслеживания рук 3, обнаружения лица 4 подключены к модулю обработки 5. К каждому из упомянутых модулей 2-4 подключены счетчики 6, выполненные с возможностью фиксации последовательности и длительности каждого из элементов жеста, захваченного с помощью модуля захвата жеста 1.

Выход модуля обработки 5 подключен к модулю сравнения 7. К одному из выходов модуля сравнения 7 подключен блок матриц кирем 8. К другому из выходов модуля сравнения 7 последовательно подключены модуль накопления 9, модуль идентификации 10, блок матрицы жестов 11, интерфейсный модуль 13 и блок ввода данных 14.

К модулю идентификации 10 параллельно подключен генератор чисел 12. К блоку матриц кирем 8 параллельно подключен блок матриц жестов 11. Интерфейсный модуль 13 подключен к модулю захвата жеста 1.

Способ генерации и распознавания жестов осуществляют следующим образом.

Формируют массив данных жестов. Для этого с помощью профессионального диктора или пользователя, профессионально владеющего жестовым языком, воспроизводят жесты жестового языка.

С помощью модуля захвата 1 захватывают в режиме реального времени эталонное изображение жеста, воспроизводимое, как указано выше, диктором или профессиональным пользователем, записывают его и известными способами осуществляют обнаружение объектов на захваченном изображении с помощью модуля обнаружения объектов 2. Для обнаружения объектов используют специальное программное обеспечение, например, Mediapipe. При обнаружении объектов модуль обнаружения объектов 2 передает в счетчик 6 соответствующее выявленному количество объектов.

После обнаружения объектов в модуле обнаружения объектов 2, в модуле отслеживания рук 3 и модуле обнаружения лица 4 по маркерным точкам, отмеченным на скелете человека, осуществляют распознавание наличия человека в кадре, его позы, распознавание рук и каждого из пальцев кистей рук, их местоположение и расположение относительно друг друга, тела, движение рук, локтей, распознавание ключевых точек мимики лица пользователя. С помощью счетчиков 6 осуществляют подсчет в каждом

из упомянутых модулей 3-4 количества обнаруженных точек и их динамических изменений. Обнаруженные и распознанные элементы жеста передают в модуль обработки 5, в котором осуществляют его разделение на киремы по характерным точкам для частей тела пользователя, подсчет кирем с помощью счетчика 6, их последовательности на шкале времени и формируют цифровую графическую модель жеста.

Выделенные в модуле обработки 5 киремы передают в модуль сравнения 7, в котором осуществляют их сравнение с киремами, уже сгенерированными и хранящимися в блоке матрицы кирем 8, то есть идентификацию кирем, присвоение им соответствующих аргументов и значений этих аргументов.

При успешной идентификации киремы данные о ней передают в модуль накопления 9, где по аргументам, их значениям и их последовательности формируют формулу жеста для использования ее в дальнейшем для цифровой графической модели (аватара).

При отсутствии в блоке матрицы кирем 8 соответствующей распознаваемой киремы через блок ввода данных 14 осуществляют запись киремы с присвоением вручную в блоке матрицы кирем 8 текстового наименования и описания киремы, присваивают с помощью генератора чисел 12 значение аргументам кирем.

После формирования формулы жеста с помощью блока ввода данных 14 в блок матрицы жестов 11 вводят текстовое наименование распознанного жеста, его описание и подгружают из блока матриц кирем 8 описание его кирем, включающее их конфигурации (аргументы и значения этих аргументов, загружают изображение с реперными точками жеста, видео жеста. С помощью генератора чисел 12 присваивают идентификатор жесту.

После формирования массива (базы) данных жестов и кирем и обучения цифровых графических моделей, модуль распознавания 1 работает без специального оборудования и специального помещения, используя только одну камеру (веб-камеру).

Под цифровой графической моделью в настоящем изобретении понимают цифровой двойник человека, на вход которого подают цифровую графическую модель жеста в виде формулы жеста, а на выходе получают графическую модель человека, воспроизводящую по выделенным характерным точкам для частей тела пользователя киремы жеста в принятой для жестовой вселенной системе координат и жест в целом. Подача на вход цифрового двойника человека формулы жеста с целью получения на выходе графической модели человека, воспроизводящей жест по его формуле, называют генерацией жестов.

Система координат в простейшем варианте реализации представляет собой систему, где каждая точка (пиксель) на общей цифровой графической модели имеет три измерения x , y и z , которые характеризуют ее положение в пространстве модели, т.е. каждая точка характеризуется набором параметров $[x, y, z]$, например координаты точки $A = [10 \text{ px}, 350 \text{ px}, 78 \text{ px}]$. Параметр может быть привязан как к точке, так и к области жестовой вселенной.

Изображение общей цифровой графической модели на экране интерфейсного модуля 13 устройства для распознавания жестов может быть реализовано с использованием Microsoft Silverlight, Adobe Flash, HTML 5 и других технологий.

Алгоритм распознавания жестов показан на Фиг. 7, который работает следующим образом.

С помощью модуля захвата 1 захватывают в режиме реального времени изображение обычного пользователя, воспроизводящего жест и осуществляют обнаружение объектов на захваченном изображении с помощью модуля обнаружения объектов 2. Для

обнаружения объектов используют специальное программное обеспечение, например, Mediapipe. При обнаружении объектов модуль обнаружения объектов 2 передает в счетчик 6 соответствующее выявленному количеству объектов.

После обнаружения объектов в модуле обнаружения объектов 2, в модуле
5 отслеживания рук 3 и модуле обнаружения лица 4 по маркерным точкам, отмеченным на скелете человека, осуществляют распознавание наличия человека в кадре, его позы, распознавание рук и каждого из пальцев кистей рук, их местоположение и расположение относительно друг друга и тела, движение рук, локтей, тела, распознавание ключевых точек мимики лица пользователя, интенсивности и амплитуды выполнения жеста,
10 фиксацию перемещения или статичности для неперемещаемого жеста.

С помощью счетчиков 6 осуществляют подсчет в каждом из упомянутых модулей 3-4 количества обнаруженных точек и их динамических изменений. Обнаруженные и распознанные элементы жеста передают в модуль обработки 5, в котором осуществляют его разделение на киремы по характерным точкам для частей тела пользователя, подсчет
15 кирем с помощью счетчика 6, их последовательности на шкале времени.

Выделенные в модуле обработки 5 киремы передают в модуль сравнения 7, в котором осуществляют их сравнение с киремами, уже хранящимися в блоке матрицы кирем 8, то есть идентификацию кирем, присвоение им соответствующих аргументов и значений этих аргументов. При успешной идентификации киремы данные о ней передают в
20 модуль накопления 9, где по аргументам, их значениям и их последовательности формируют формулу жеста. При отсутствии в блоке матрицы кирем 8 соответствующей распознаваемой киремы с помощью блока ввода данных 14 определяют (уточняют) значение этой киремы и осуществляют ее кирематический разбор с учетом уже известных и включенных в блок матриц кирем 8.

Если показываемая кирема является новой, то ее включают в массив данных в качестве эталонной при условии ее разбора и воспроизведения профессиональным диктором или пользователем, профессионально владеющим жестовым языком, присваивают ей соответствующий аргумент и новое значение этого аргумента.

Если показываемая кирема является не новой, но вариантом одной из существующих,
30 характеризующей, например, почерк пользователя, то ее включают в массив данных в качестве варианта существующей киремы (люфта или дельты киремы), с присвоением ей того же самого аргумента и его значения.

Если показываемая кирема по результатам кирематического разбора жеста признана ошибочной в показе, корректируют эту кирему в жесте пользователя и повторно
35 воспроизводят и захватывают ее для распознавания. Люфт вводят для распознавания жестов, который нивелирует влияние личного почерка человека. Например, буква «В» и цифра «5» (см.Фиг.8). Указанные на фиг.8 жесты отличаются расстановкой пальцев. Дельта определяется разницей координат X, Y, Z.

Наличие формулы жеста позволяет решить обратную задачу, которая позволяет из
40 формулы получить словесное описание жеста. Пример реализации: человек владеющий жестовым языком показывает жест системе. Система распознает этот жест и воспроизводит его словесное или голосовое значение.

Технический результат изобретения заключается в повышении точности распознавания жеста жестового языка, в том числе с учетом индивидуального почерка
45 жестикулирующего за счет увеличения количества детектируемых элементов (кирем), создания с помощью заявленного устройства базы данных эталонных элементов и жестов и дельт отклонения исполнения этих элементов, присущих индивидуальному почерку жестикулирующего, присвоения каждому из элементов уникального

обозначения и формирования из кирем с учетом их последовательности, распределения во времени и во взаимном временном расположении относительно друг другу в жесте (таймлайн) в виде реляционной уникальной формулы для использования в последующем в цифровой графической модели человека в качестве эталона для распознавания. При этом вновь распознаваемые киремы не исключают из рассмотрения при отрицательной идентификации в отличие от существующего уровня техники, а уточняют ее значение и если кирема является новой, то ее включают в матрицу кирем в качестве эталонной после кирематического разбора и присвоения ей уникального обозначения, если показываемая кирема является вариантом одной из существующих в матрице кирем, то ее включают в матрицу кирем в качестве варианта существующей киремы, с присвоением ей уникального обозначения существующей киремы, если показываемая кирема по результатам кирематического разбора жеста признана ошибочной в показе, корректируют эту кирему в жесте и повторно воспроизводят и захватывают для распознавания. Таким образом, заявленное изобретение способствует самообновлению и расширению массива данных, что в результате с расширением массива данных повышает вероятность и точность распознавания жеста.

(57) Формула изобретения

1. Способ генерации и распознавания жестов жестового языка, характеризующийся тем, что первоначально формируют массив данных содержащий матрицу жестов, включающую в себя идентификатор, текстовое наименование, описание жеста, изображение жеста и связанную с матрицей жестов матрицу кирем жестов, включающую киремы и их конфигурации, для чего при воспроизведении жеста осуществляют его захват, разделение на киремы, из которых формируют матрицу кирем, каждой из кирем присваивают уникальное обозначение в матрице по группе, к которой относится кирема, и списку возможных конфигураций этой киремы в этой группе для последующей идентификации киремы в матрице кирем и распознавания жеста в связанной с матрицей кирем матрице жестов, при этом при разделении жеста на киремы учитывают их последовательность, количество, распределение во времени и во взаимном временном расположении относительно друг друга в жесте и представляют в виде реляционной формулы, уникальной для каждого жеста и содержащей последовательность, количество, таймлайн кирем и списочное значение каждой из кирем в ее группе для последующей генерации жеста путем подачи формулы жеста на вход графической модели человека и воспроизведения этой графической моделью человека жеста в принятой для жестовой вселенной системе координат, для распознавания жеста захватывают изображение жеста, осуществляют разделение жеста на киремы, которые сравнивают с киремами в матрице кирем и осуществляют идентификацию кирем, при этом при отсутствии в матрице кирем уточняют значение этой киремы, и если кирема является новой, то ее включают в матрицу кирем в качестве эталонной после кирематического разбора и присвоения ей уникального обозначения, если показываемая кирема является вариантом одной из существующих в матрице кирем, то ее включают в матрицу кирем в качестве варианта существующей киремы с присвоением ей уникального обозначения существующей киремы, если показываемая кирема по результатам кирематического разбора жеста признана ошибочной в показе, корректируют эту кирему в жесте и повторно воспроизводят и захватывают для распознавания.

2. Способ по п. 1, отличающийся тем, что жест разбивают на киремы, характеризующие количество рук, используемых в жесте, конфигурацию ведущей руки, конфигурацию вспомогательной руки, положение кистей относительно тела, положение

кистей относительно друг друга, движение кистей, движение локтей, движение тела, движение рук, область выполнения жеста, мимику лица, интенсивность и амплитуду выполнения жеста, перемещаемость жеста.

3. Способ по п. 1, отличающийся тем, что конфигурации вспомогательной руки являются зеркальным отображением конфигураций ведущей руки.

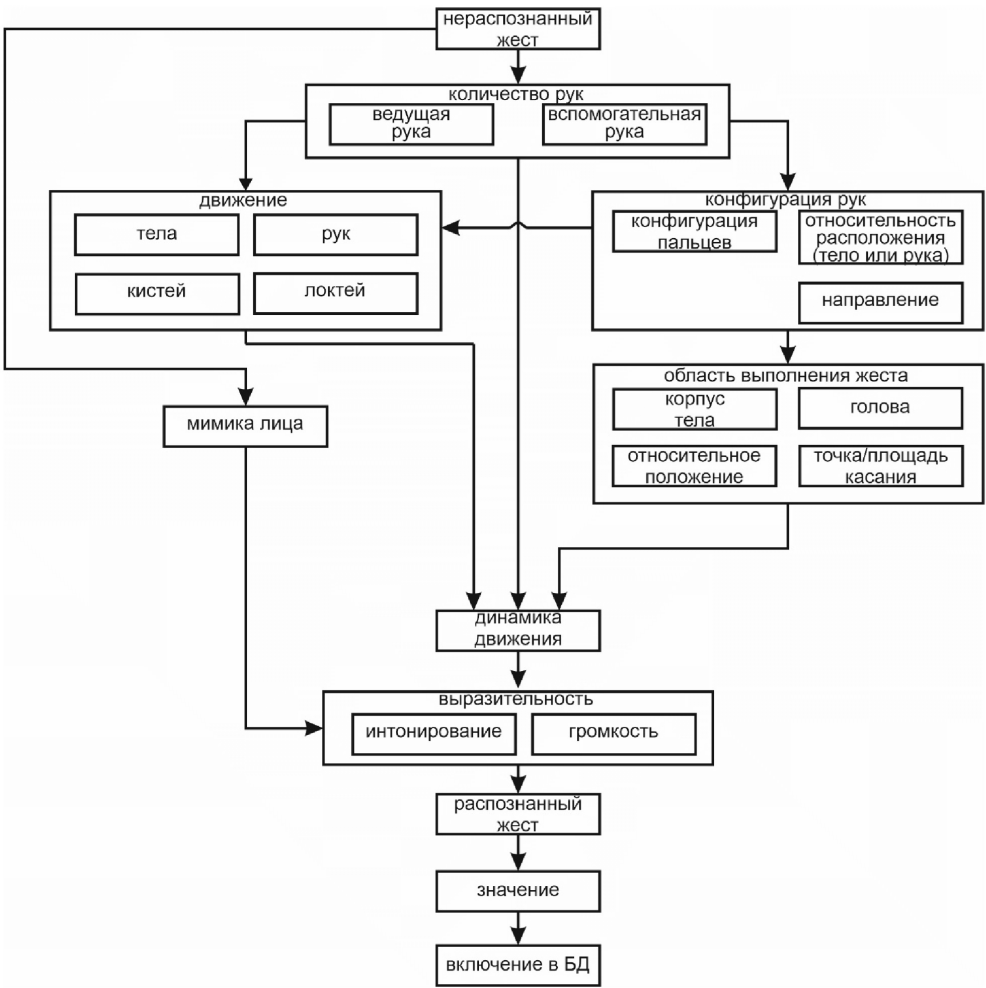
4. Способ по п. 1, отличающийся тем, что количество кирем в жесте в процессе изучения или развития жестового языка может увеличиваться.

5. Способ по п. 1, отличающийся тем, что для включения жеста в массив данных жестов в качестве эталонного жест воспроизводит профессиональный диктор или пользователь, профессионально владеющий жестовым языком.

6. Устройство для генерации и распознавания жестов, включающее модуль захвата жеста, к модулю захвата жеста последовательно подключены модуль обнаружения объектов на захватываемом с помощью модуля захвата жеста изображении, модуль отслеживания рук и модуль обнаружения лица, модули обнаружения объектов, отслеживания рук, обнаружения лица подключены к модулю обработки, при этом к каждому из модулей обнаружения объектов, отслеживания рук, обнаружения лица подключены счетчики, выполненные с возможностью фиксации последовательности и длительности каждого из элементов жеста, захваченного с помощью модуля захвата жеста для формирования формулы жеста, выход модуля обработки подключен к модулю сравнения, к одному выходу модуля сравнения подключен блок матриц кирем, к другому из выходов модуля сравнения последовательно подключены модуль накопления, модуль идентификации, блок матрицы жестов, интерфейсный модуль и блок ввода данных, к модулю идентификации параллельно подключен генератор чисел, к блоку матриц кирем параллельно подключен блок матриц жестов, интерфейсный модуль подключен к модулю захвата жеста, при этом интерфейсный модуль выполнен с возможностью отображения цифровой графической модели человека, воспроизводящей жест по формуле жеста, полученной при распознавании жеста.

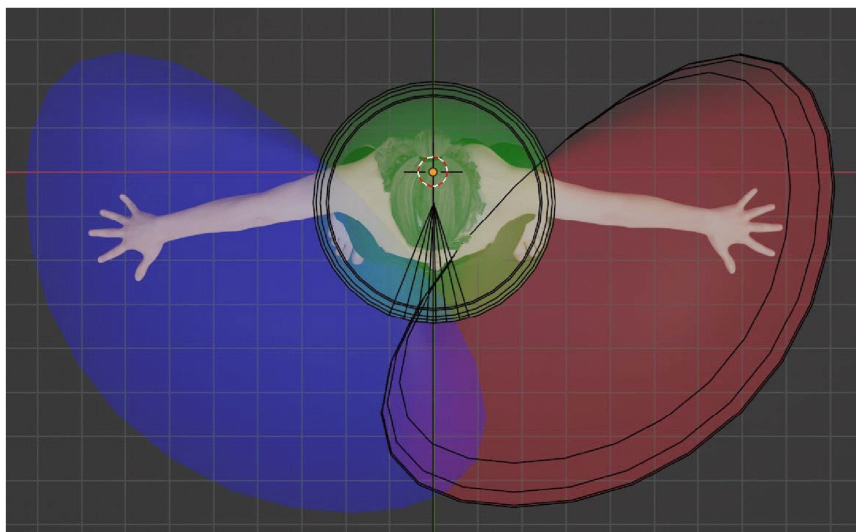
7. Устройство по п. 6, отличающееся тем, что модуль захвата жеста выполнен в виде устройств машинного зрения, например стереокамер или систем камер и датчиков глубин выполненных с возможностью получать X, Y, Z - координаты жеста.

1

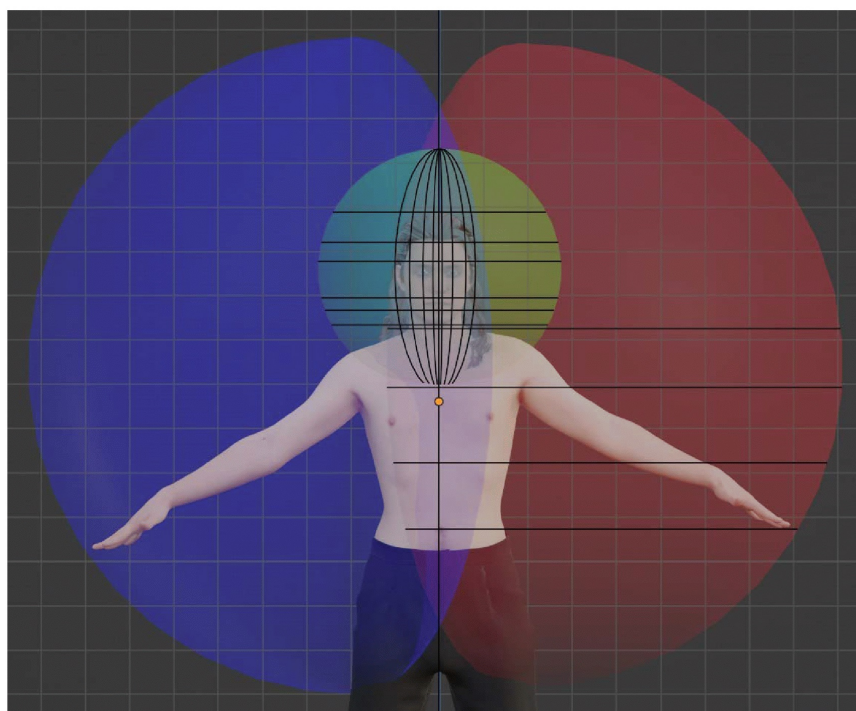


Фиг.1

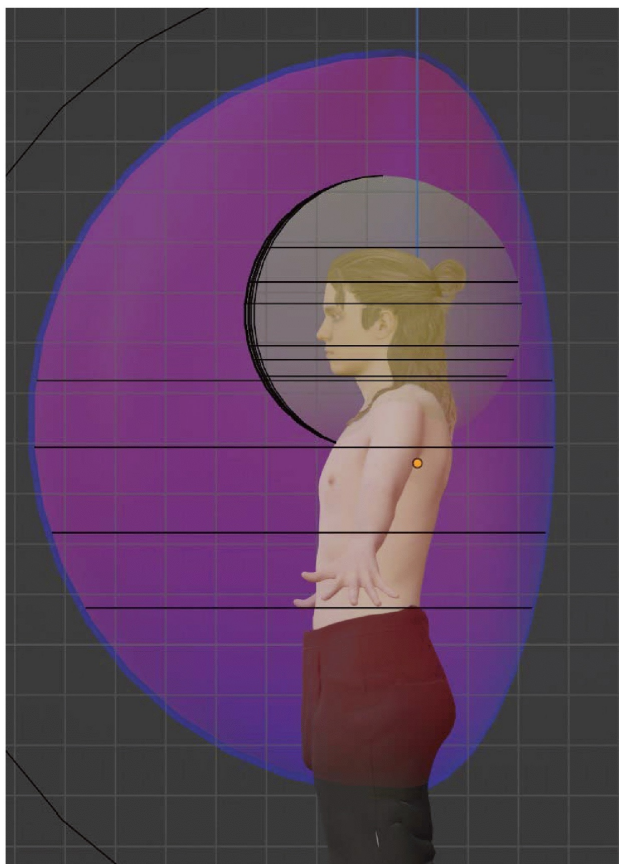
2



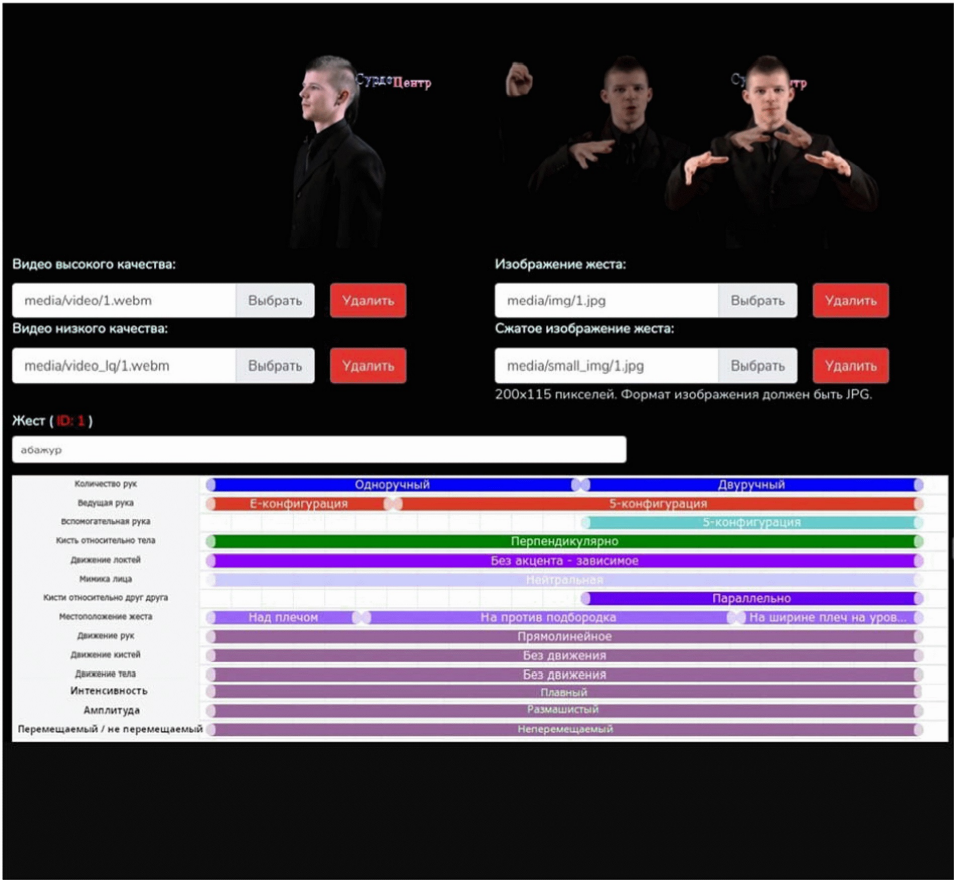
Фиг.2



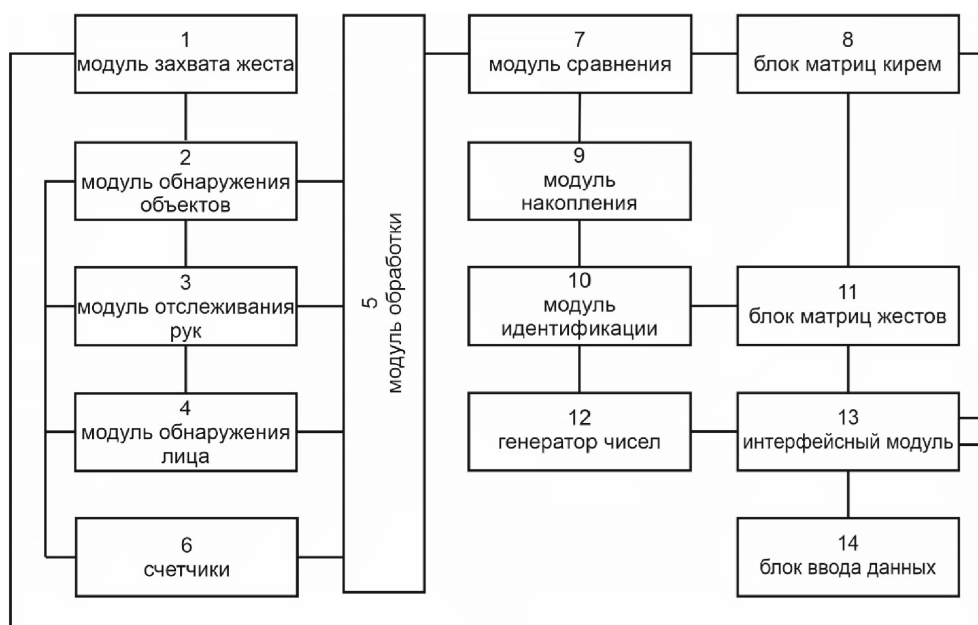
Фиг.3



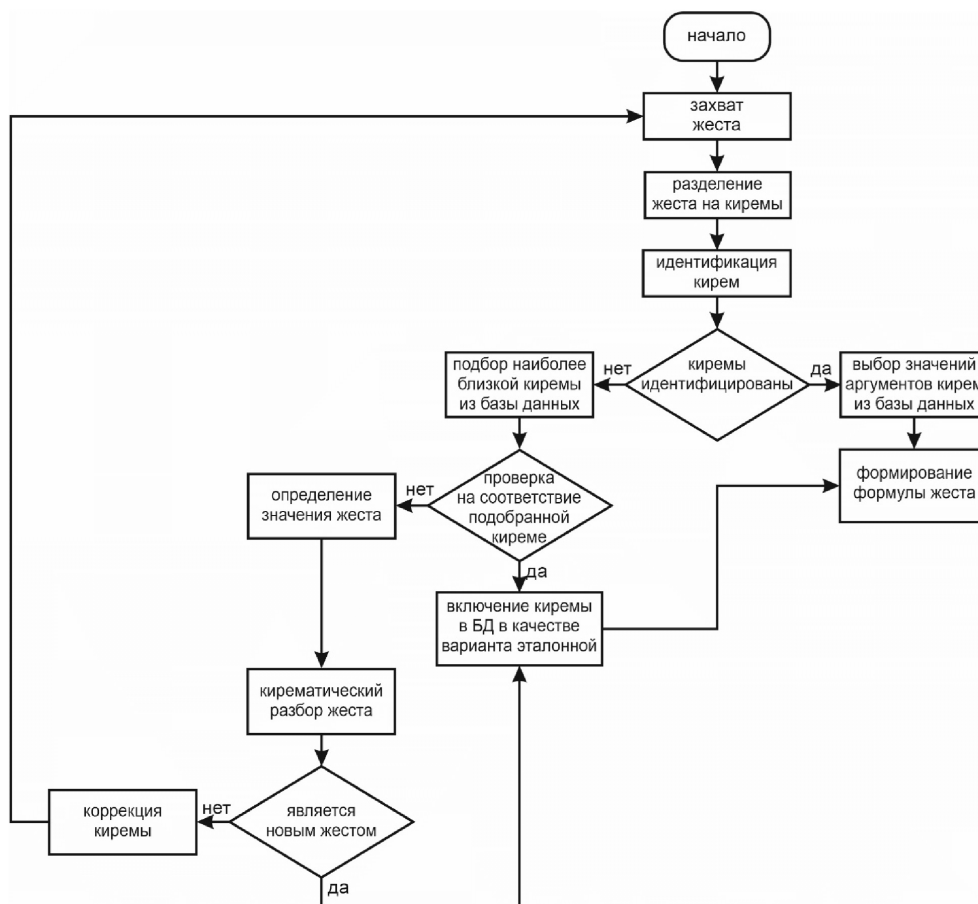
Фиг.4



Фиг.5



Фиг.6



Фиг.7



Фиг.8