

В. А. Зенг

*Омский государственный технический университет
пр. Мира, 11, Омск, 644050, Россия*

valeriyazeng@mail.ru

ФОРМИРОВАНИЕ БАЗОВОГО СЛОВАРЯ ЖЕСТОВ ДЛЯ ЕСТЕСТВЕННОГО КОМПЬЮТЕРНОГО БЕСКОНТАКТНОГО ИНТЕРФЕЙСА *

Исследуются возможности бесконтактных систем и интерфейсов, главные принципы работы с такими технологиями. Рассмотрена возможность применения подобных систем для упрощения взаимодействия пользователей с ограничениями возможностями здоровья с компьютерным интерфейсом. Приведены особенности и преимущества использования естественных интерфейсов и систем, основанных на жестовом управлении. Также детально рассмотрены этапы формирования базового словаря жестов для дальнейшего его применения в бесконтактном интерфейсе. В качестве дополнительного аппаратного обеспечения для получения более точных результатов распознавания таких жестов рассмотрено устройство Microsoft Kinect.

Ключевые слова: человеко-машинное взаимодействие, программное приложение, компьютерный интерфейс, прототип программы, захват движения, технологии бесконтактного взаимодействия, пользователи с ограниченными возможностями здоровья.

Введение

Современные научные работы, посвященные исследованию человеко-машинного взаимодействия, направлены в основном на создание вычислительных машин, оборудованных большим количеством различных датчиков и сенсоров, а также на изучение средств межлической коммуникации, таких как речь и сопровождающие жесты. Разрабатываемые интерфейсы ориентированы исключительно на опытных пользователей, однако почти не затрагиваются вопросы человеко-машинной коммуникации для лиц с ограниченными возможностями (инвалидов). Так, глухонемые люди не могут использовать речевые интерфейсы, а люди с проблемами мелкой моторики не способны работать с клавиатурой или жестовыми интерфейсами. Главной целью является разработка универсального бесконтактного интерфейса, пригодного для всех категорий пользователей, и реализация этого интерфейса, демонстрирующего возможности многомодальной человеко-машинной коммуникации. Такой интерфейс будет включать различные естественные для человека способы передачи и восприятия информации: речь, жесты, движения головой и телом, чтение по губам, а также комбинации этих бесконтактных модальностей. Многомодальный интерфейс сможет обрабатывать входную информацию и выводить информацию в той форме, которая доступна конкретному пользователю [1].

* Работа выполнена при поддержке Фонда содействия инновациям в рамках программы «УМНИК», договор № 12394ГУ/2017.

Возможности бесконтактных интерфейсов

Естественный интерфейс подразумевает, что основным путем взаимодействия человека с компьютером являются прикосновения, жесты, речь, а также другие виды поведения, которые практикуются в течение многих лет и/или которые являются врожденными. Прежде всего, понятие естественного интерфейса относится к процессу взаимодействия пользователя с системой – насколько оно комфортно и понятно. Управление должно быть одинаково простым, интуитивным, как для опытного пользователя системы, так и для новичка [2].

Пользователи предпочитают избавляться от барьеров, воздвигаемых традиционными интерфейсами между ними и техникой, – гораздо удобнее управлять устройствами без посредников, а жесты – вполне естественный способ общения. Интерфейсы на жестыкуляционном управлении предлагают привычными движениями, например, имитирующими перелистывание страницы, реализовать то же самое действие на экране. Нажатие на клавиши не относится к естественным способам общения для человека, поэтому основное преимущество жестыкуляционного управления в том, что оно позволяет быстрее и проще отдавать команды устройствам. Помимо этого, подобные интерфейсы позволяют избавиться экранное пространство от кнопок, клавиш и других наглядных элементов управления [3]. Кроме того, жестыкуляция дает возможность сосредоточиться на экране, а не на компьютерной мыши, клавиатуре или пульте дистанционного управления.

Для реализации необходимо разработать интерфейс для работы с компьютером при помощи жестов, предварительно задаваемых пользователем. Подобный интерфейс предназначен, в основном, для помощи следующим категориям пользователей:

- людям с проблемами мелкой моторики;
- слабослышащим или глухонемым людям.

Жестыкуляционные интерфейсы могут упростить инвалидам взаимодействие с электроникой. Помимо этого, существует еще несколько областей, где возможно их потенциальное применение. Одна из них – автомобили. С помощью телодвижений можно управлять развлекательной системой, очистителями стекол, фарами и другим оборудованием, не отрывая глаз от дороги. «Ford Motor» уже выпускает автомобили, у которых автоматически открывается багажник, если под ним провести ногой [1].

Жестыкуляционные интерфейсы также позволят врачам и медсестрам управлять компьютерами и другими устройствами, не дотрагиваясь до них. Это очень ценная возможность в ситуациях, когда медикам нужны чистые руки, а также в случаях, когда оборудование находится вне прямой досягаемости.

Пользователь жестовой системы управления должен в первую очередь догадаться о следующем:

- что системой можно управлять с помощью жестов;
- какой набор жестов поддерживает эта система.

Эту информацию пользователь должен узнать при работе с самой системой, а не из руководства или другой документации. В целом довольно трудно ответить на вопрос, как достигнуть «affordance» (с англ. – воспринимаемая доступность) управления жестами. Термин «affordance» означает это качество системы или продукта, которое предложил Джеймс Гибсон, основатель направления в психологии, рассматривающего восприятие как процесс, не предполагающий умозаключений, промежуточных переменных. В данной ситуации, обратную связь целесообразно дополнить подсказками-предсказаниями – в процессе распознавания показать, какие жесты доступны на данном этапе. Помимо конечной цели, пользователь сможет узнать, какие еще возможности поддерживает система. Также, если у некоторых жестов траектория одинаково начинается, это может стать полезным для пользователя дополнением [4].

Особенности систем жестового управления

Управление интерфейсом должно быть интуитивным и простым, он должен легко настраиваться и работать. В частности, нужно помнить, что пользователи не должны знать, как устроена система и как она работает, поэтому ошибочное распознавание и неверное испол-

зование в идеале необходимо выявлять на стадии тестирования и игнорировать его. В случае если система неверно распознала жест, необходимо предусмотреть простой механизм отмены действия и возможность отметить, что жест распознан неправильно [5]. Чтобы персонифицировать набор жестов, необходим интерфейс, который предполагает определение персональных жестов.

Понятие доступности требует «равных прав для людей в получении доступа к информации независимо от физических и когнитивных затруднений, которые они могут испытывать в связи с временными или хроническими нарушениями и болезнями». Управление жестами должно быть доступно для пользователей с разными физическими или техническими ограничениями [6]. Также в идеале при управлении в жестовых системах не должны накладываться ограничения на окружающую среду, в которой они используются. Фактически система должна быть работоспособна в любое время, при любом освещении и при любом пространственном положении¹. На сегодняшний день распознавание жестов в таких интерфейсах базируется преимущественно на получении данных с одной или нескольких камер, а это накладывает ограничения на среду, в которой они используются. Помимо этого, калибровка камеры может оказаться достаточно сложной задачей для среднестатистического пользователя.

При разработке системы жестового управления есть два варианта: создать универсальный набор жестов или дать возможность пользователям самостоятельно определять этот набор [5]. В первом случае достаточно сложно персонифицировать жесты, так как многое может зависеть от культуры пользователей, их личных особенностей (правша или левша) и других характеристик, которые влияют на естественные жесты пользователей. Поэтому интерфейсу, основанному на жестовом управлении, требуется либо приспособляться к персональным жестам, либо модифицировать универсальные жесты, либо определять персональные жесты. Возможность приспособляться к персональным жестам, основанная на наблюдении, предполагает обнаружение оптимальных жестов для такой задачи, однако это довольно длительный процесс, следовательно, этот вариант может не подойти для ежедневного использования [4]. Остальные варианты предпочтительнее всего для повседневных систем. Для их реализации нужно применять алгоритмы машинного обучения. При автоматической сегментации могут возникнуть проблемы алгоритма распознавания, так как захватываются движения, которые не являются частью жеста. Подобные движения считаются шумом, поэтому машинные модели обучения это должны учитывать².

В целом список жестов для управления компьютером должен представлять собой какой-то базовый набор жестов, и при этом некоторые из них могут быть уникальными в рамках всей системы (например, жест отмены), а некоторые – не уникальны, так как один и тот же жест в зависимости от контекста может использоваться по-разному [5]. Однако при этом у системы должна быть возможность адаптации (персонализации) под каждого конкретного пользователя, а также под условия использования интерфейса.

Разработка базового словаря жестов

Формирование набора естественных жестов для бесконтактного управления имеет корни в ставших обыденными жестах при сенсорном взаимодействии, используемом в современной цифровой технике: смартфонах, планшетах, некоторых ноутбуках (laptop). Интерактивная технология «multi-touch interaction» (MT) позволяет пользователю путем прикосновений управлять графическим интерфейсом одновременно несколькими пальцами руки.

Поддерживаемые жесты: «щелчок (выбор элемента)», «сдвиг (пролистывание) вправо», «сдвиг влево», «уменьшение масштаба», «увеличение масштаба», «поворот» [2] (рис. 1). Данный набор жестов стал основой для базовых жестов бесконтактного взаимодействия

¹ ГОСТ Р ИСО 20282-1-2011 «Эргономика изделий повседневного использования. Часть 1. Требования к конструкции элементов управления с учетом условий использования и характеристик пользователя». Введен 01.12.2012. М.: Стандартинформ, 2012. 24 с.

² ГОСТ Р ИСО 9241-20-2014 «Эргономика взаимодействия человек-система. Часть 20. Руководство по доступности оборудования и услуг в области информационно-коммуникационных технологий». Введен 01.12.2015. М.: Стандартинформ, 2015. 44 с.

с той лишь разницей, что главными управляющими «инструментами» стали не пальцы, а кисти и руки. Исходя из этого «масштаб» управления был расширен от экрана сенсорного телефона до бесконтактного управления интерфейсом с помощью рук.

Прежде всего следует определить условные обозначения, которые будут указаны в словаре жестов как начальное, конечное положение руки и направление ее движения. Для этого используются интуитивно понятные графические примитивы, которые дополняются сопроводительными подписями (рис. 2).

Основными движениями в данном бесконтактном интерфейсе являются перелистывание вправо и влево, а также прокрутка вверх и вниз. Это базовые движения, которые на сегодняшний день интуитивно понятны любому пользователю, который хотя бы раз работал с персональным компьютером. Эти жесты осуществляются взмахом руки в нужную сторону (рис. 3). Они необходимы в ситуациях, когда информация полностью не помещается на экране монитора, разделяется на части (страницы), поэтому для просмотра новой порции информации требуется переход в следующую часть.

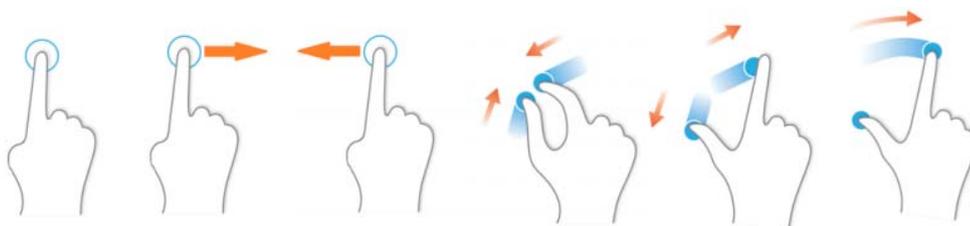


Рис. 1. Набор стандартных жестов для управления цифровым сенсорным устройством



Рис. 2. Условные обозначения словаря жестов и их расшифровка

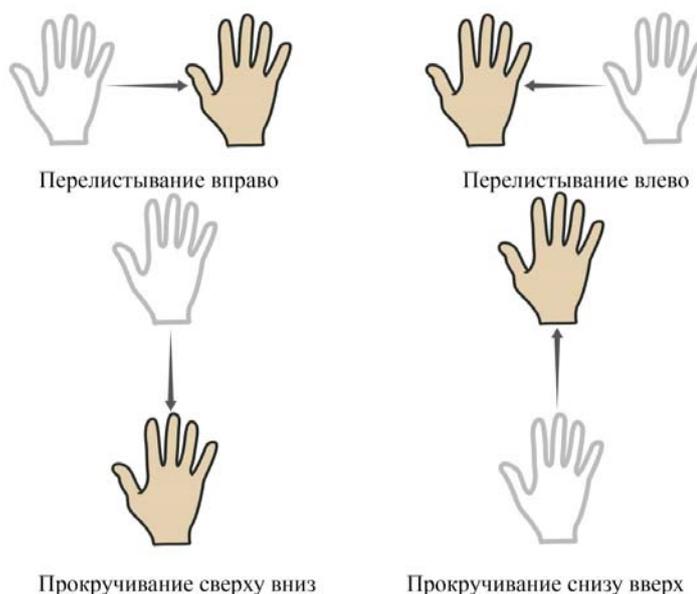


Рис. 3. Жесты пролистывания вправо или влево, прокручивания вверх и вниз



Рис. 4. Жесты предпросмотра содержимого папки и просмотра панели «Пуск»

Для облегчения взаимодействия пользователя с ограниченными возможностями здоровья (ОВЗ) с интерфейсом было разработано такое функциональное улучшение, как «предпросмотр содержимого». Зачастую пользователи, которые сортируют документы по множеству папок, не могут вспомнить точное месторасположение того или иного файла. Поэтому была создана функция «предпросмотра содержимого», которая позволяет вывести список последних открытых в определенных папках файлов, что может сэкономить много времени на поиск необходимых документов. Данная возможность реализуется путем изображения круга по часовой стрелке для раскрытия списка файлов и против часовой для их закрытия (рис. 4).

Также было принято решение создать новое движение, которое упростит работу с меню, подобным элементу «Пуск» в ОС Windows. Так как у пользователей с ограниченной моторикой возникает необходимость совершать действия в системе, затрачивая как можно меньше сил и совершая не очень активные движения, то было принято решение сделать жест максимально понятным и удобным для людей с ОВЗ [1]. Для того чтобы отобразилась панель меню, следует немного поднять ладони вверх, как при прокручивании снизу. Чтобы панель скрыть, следует сделать такое же движение, но в обратном направлении (рис. 5). Однако многие пользователи с ограничениями в движении не имеют возможности двигать две руки одновременно, к тому же в одном направлении. Поэтому интерфейс предоставляет возможность совершать данные манипуляции одной рукой.

Чтобы совершить «клик» по выбранному жестом элементу интерфейса, следует удерживать ладонь на месте 3 секунды. Если задать возможную амплитуду размаха руки в попытках удержать ее на одном месте, что не всегда удается людям с ограниченной моторикой из-за судорог или непроизвольных спазмов мышц, то можно настроить систему для восприятия дрожащей или даже колеблющейся в пределах 5–10 см руки как жест выбора элемента [3].

Последней из возможных базовых функций, адаптированных под пользователей с ОВЗ, является функция увеличения и уменьшения размеров объектов. При просмотре изображения или документа в его начальном состоянии не всегда достигаются необходимые для корректного восприятия величины объекта с самого его запуска. Поэтому функция увеличения и уменьшения размеров особенно необходима. Увеличение достигается путем раздвигания рук из центра в стороны по диагонали (см. рис. 5). Не имеет значения, какая именно рука окажется в верхней точке жеста, а какая в нижней. Система распознает оба положения как

один и тот же запрос. Соответственно уменьшение элемента может быть получено в том случае, когда руки, разведенные в стороны и находящиеся по диагонали друг от друга, собираются в центре. Причем степень увеличения и уменьшения может достигаться как поэтапно, если указать в настройках, что одно такое движение рук – это один шаг масштабирования элемента, так и плавно, что актуально для пользователей без ограничений в моторике.

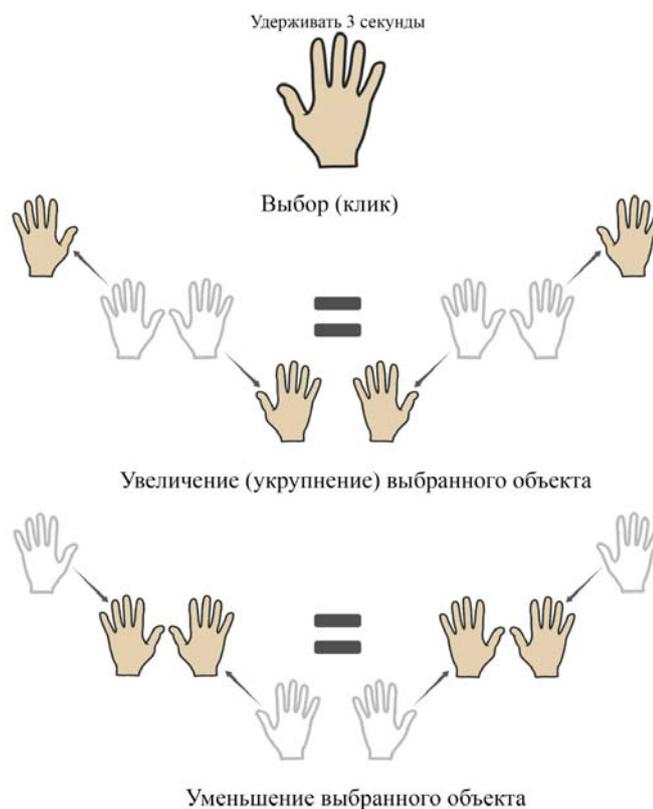


Рис. 5. Жесты выбора элемента и увеличения или уменьшения объекта

Таким образом, был разработан базовый словарь жестов для естественного бесконтактного взаимодействия с компьютерным интерфейсом, который удовлетворяет основным принципам естественных интерфейсных систем.

- Быстрое погружение. Прилагая минимальные усилия и время, человек сам овладевает системой. Отсюда не следует, правда, что обучения вообще не должно быть, но оно должно быть минимальным и без посторонней помощи. Важно понимать, что мы говорим именно об интерфейсе, естественный интерфейс может быть и у промышленного робота, и у баллистической ракеты, и у шахмат, но овладение им в данных контекстах не сделает вас опытным пользователем с точки зрения системы. Недостаточно понять базовые элементы и правила их комбинирования, нужно изучить их идиоматическое использование.

- Легкое управление. Компьютер должен приспосабливаться к пользователю, а не наоборот. Человек через некоторое время перестает «думать» о том, что ему надо сделать, чтобы произошло нужное действие.

- Впечатление (игровой момент). Эффект новизны пользовательского опыта, как следствие, положительные эмоции от самой системы. Со временем и частотой использования этот эффект будет исчезать [4].

Рынок систем жестикуляционного управления относительно нов, стандартов еще практически нет, и в различных системах используются совершенно разные интерфейсы, камеры и алгоритмы. Это дает возможность выбора инструментов для использования в разработке и применении естественных интерфейсов. В разрабатываемом продукте предполагается использование устройства Microsoft Kinect в качестве дополнительного аппаратного обеспече-

ния. Оно определяет до шести человек в пространстве и 25 суставов каждого из них; распознает лица, эмоции, пульс. Аудиосистема последней модели может определить двух одновременно говорящих людей и распознать два потока речи [3]. Помимо широких возможностей и относительно низкой цены по сравнению с аналогами, устройство Microsoft Kinect может также адаптироваться под особенности пользователей, просчитывая возможные варианты жестов. Так что если человек с ОВЗ не повторит определенное движение в точности, как указано в словаре жестов, то устройство захвата движения все равно идентифицирует кисть, распознает движение и определит его тип в соответствии со словарем.

Выводы

Таким образом, создание и внедрение естественных человеко-машинных интерфейсов, основанных на автоматическом распознавании речи и жестов, предлагает пользователям-инвалидам новый способ персонифицированного бесконтактного взаимодействия с компьютером полностью без использования стандартных устройств ввода, таких как клавиатура и компьютерная мышь. Адаптивность и индивидуализация интерфейса, а также способность осуществить настройку словаря движений для управления системой в соответствии с потребностями каждой категории граждан с ОВЗ являются отличительной особенностью разработки. Еще одним важным дополнением будет возможность встраивания его в виде программного модуля для управления операционной системой MS Windows и ее системными приложениями.

Список литературы

1. Ронжин А. Л. Методы и многомодальные интерфейсы для бесконтактной коммуникации инвалидов с информационно-справочными системами / Российский фонд фундаментальных исследований. М., 2016. URL: http://www.rfbr.ru/rffi/portal/project_search/o_44865 (дата обращения 11.05.2018).
2. Дроздова Ю. А. Создание естественного пользовательского интерфейса для мобильных устройств. Обзор средств распознавания жестов руки // Студенческий научный форум: Материалы VI Междунар. студ. электрон. науч. конф. 2014. URL: <http://www.scienceforum.ru/2014/527/1470> (дата обращения 11.05.2018).
3. Зенг В. А. Анализ подходов к созданию бесконтактных интерфейсов // Творчество молодых: дизайн, реклама, информационные технологии: Материалы XV Междунар. науч.-практ. конф. Омск, 2016. С. 107–111.
4. Гарбер Л. Новые пользовательские интерфейсы: жестикуляция // IEEE Computer Society, Synapses to Circuitry: Gestural Technology: Moving Interfaces in a New Direction. 2013. № 10. URL: <https://www.osp.ru/os/2013/10/13039069/> (дата обращения 11.05.2018).
5. Сорокин Д. Упрощение в дизайне интерфейсов. URL: <https://www.uplab.ru/blog/simplification-in-the-design-of-interfaces/> (дата обращения 11.05.2018).
6. Николахина Е. Равные права – равные возможности // Библиотека. 2015. № 10. С. 52–54.

Материал поступил в редколлегию 11.05.2018

V. A. Zeng

*Omsk State Technical University
11 Mir Ave., Omsk, 644050, Russian Federation*

valeriyazeng@mail.ru

GENERAL GESTURAL DICTIONARY DEVELOPMENT FOR NATURAL COMPUTER-BASED CONTACTLESS INTERFACE

This article contains the description of contactless systems and interfaces and main principles of working with these technologies. There is possibility of using such systems to simplify the interac-

tion of users with the limitations of health possibilities with a computer interface. The features and advantages of using natural interfaces and systems based on gesture control are also presented. There are stages of the formation of a basic gesture dictionary for further use in contactless interface which are described in detail as well. As additional hardware for obtaining more accurate results of recognizing such hand gestures a Microsoft Kinect device was reviewed.

Keywords: human-machine interaction, software application, computer interface, program prototype, motion capture, contactless interaction technologies, users with disabilities.

References

1. Ronzhin A. Methods and multimodal interfaces for contactless advertising. Moscow, Russian Foundation for Basic Research, 2016. URL: http://www.rfbr.ru/rffi/portal/project_search/o_44865 (accessed: 05.11.2018) (in Russ.)
2. Drozdova Yu. Creating a natural user interface for mobile devices. A review of hand gesture recognition. *Proc. of the VI International Student Electronic Scientific Conference «Student Scientific Forum»*. URL: <http://www.scienceforum.ru/2014/527/1470> (accessed: 05.11.2018) (in Russ.)
3. Zeng V. Analysis of contactless interfaces creation approaches. *Young Creativity: Design, Advertising, Information Technologies: Proc. of the XV International Scientific and Practical Conference*. Omsk, 2016, p. 107–111. (in Russ.)
4. Garber L. New User Interfaces: Gesture. *IEEE Computer Society, Synapses to Circuitry: Gestural Technology: Moving Interfaces in a New Direction*, 2013, no. 10. URL: <https://www.osp.ru/os/2013/10/13039069/> (accessed: 05.11.2018) (in Russ.)
5. Sorokin D. Simplification in the design of interfaces. URL: <https://www.uplab.ru/blog/simplification-in-the-design-of-interfaces/> (accessed: 05.11.2018) (in Russ.)
6. Nikolahina E. Equal rights – equal opportunities. *Library*, 2015, no. 10, p. 52–54. (in Russ.)

For citation:

Zeng V. A. General Gestural Dictionary Development for Natural Computer-Based Contactless Interface. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 3, p. 105–112. (in Russ.)

DOI 10.25205/1818-7900-2018-16-3-105-112