

Федеральное агентство по образованию РФ
Государственное образовательное учреждение высшего профессионального
образования

Уральский государственный университет им. А.М.Горького

Математико-механический факультет
Кафедра информатики и процессов управления

ПРОЕКТИРОВАНИЕ ЧЕЛОВЕКО-КОМПЬЮТЕРНОГО ВЗАИМОДЕЙСТВИЯ, ОСНОВАННОГО НА ЖЕСТАХ

"Допущен к защите"

"__" _____ 2010 г.

Квалификационная работа
на степень магистра наук
по направлению
"Математика, прикладная
математика"

Флягина Татьяна
Алексеевна

Научный руководитель
Авербух Владимир
Лазаревич
заведующий сектором,
к.т.н.

Екатеринбург

2010

РЕФЕРАТ

Флягина Т. А., ПРОЕКТИРОВАНИЕ ЧЕЛОВЕКО-КОМПЬЮТЕРНОГО
ВЗАИМОДЕЙСТВИЯ, ОСНОВАННОГО НА
ЖЕСТАХ, квалификационная работа на степень магистра наук по
направлению "Математика, прикладная математика": стр. , рис. , библи. 12
назв., приложений .

На протяжении почти 40 лет, клавиатура и мышь оставались основным средством человеко-компьютерного взаимодействия. 4 года назад стали распространяться устройства с естественным управлением, основанным на жестах. Полная реализация потенциала жестовых систем задача будущего, но уже сегодня мы знаем, что такой способ взаимодействия позволит легко и интуитивно контролировать устройства вокруг нас. В связи с этим возникает потребность в проектировании набора жестов для управления системы.

В работе рассматриваются различные аспекты проектирования взаимодействия систем под управлением жестов.

СОДЕРЖАНИЕ

I. ВВЕДЕНИЕ.....	5
II. ОБЗОР СИСТЕМ ПОД УПРАВЛЕНИЕМ ЖЕСТОВ	7
Способ ввода.....	7
Сенсорный ввод.....	7
Сенсорные датчики	12
Акселерометр.....	12
Компьютерная мышь.....	13
Компьютерное зрение	14
Remote sensors.....	15
Пользовательский интерфейс.....	18
Материальный пользовательский интерфейс	20
Использование жестов в интерфейсе	22
III. ОБЗОР ЖЕСТОВ ВЗАИМОДЕЙСТВИЯ	24
Таксономия жестов человека	24
Классификация жестов человеко-компьютерного взаимодействия.....	29
Классификация жестов по функциональности.....	29
Классификация жестов по связи с речью.....	31
Манипуляционные жесты.....	33
Семафорные жесты	36
Жестикуляция	37
Язык жестов	38
Область применения жестов	39
IV. ХАРАКТЕРИСТИКА СИСТЕМ ПОД УПРАВЛЕНИЕМ ЖЕСТОВ	40
Ясная концептуальная модель	40
Соответствие.....	43
Обратная связь	44
Affordance.....	45
Возможность отмены	46
Адаптивность	47
Доступность	48

Вывод.....	49
V. ПРОЕКТИРОВАНИЕ ЖЕСТОВ В РАМКАХ USER-CENTER DESIGN	51
VI. ПРИЛОЖЕНИЕ	58
Примеры жестов для компьютерной мыши.....	58
Прототип системы управления презентацией	60
VII. ЗАКЛЮЧЕНИЕ	61
VIII. ЛИТЕРАТУРА	62

ВВЕДЕНИЕ

На протяжении почти 40 лет, клавиатура и мышь оставались основным средством человеко-компьютерного взаимодействия. Сегодня многие мобильные телефоны оснащены сенсорным экраном и акселерометром, что позволяет проектировщикам создавать *другие* пользовательские интерфейсы. Это не просто более продвинутый вариант графического пользовательского интерфейса (англ. Graphical user interface, GUI), это действительно что-то новое. В последнее время все больше специалистов в области человеко-компьютерного взаимодействия и смежных областях, говорят о сдвиге парадигмы в сторону более естественных способов взаимодействия, в сторону более естественных интерфейсов. Эти интерфейсы характеризуется прозрачностью, невидимостью взаимодействия со стороны пользователя.

Отличительной особенностью естественных интерфейсов является отсутствие глаголов, все глаголы заменяются действиями пользователя: вместо кнопки "ответить" пользователь просто подносит телефон к уху. Это новое естественное управление называется управлением с использованием специальных компьютерных жестов, это человеко-компьютерное управление, основанное на жестах. Основные элементарные операции, на которых строится традиционный графический пользовательский интерфейс это клик, двойной клик, перетаскивание, язык же управления на базе жестов более богатый, он состоит из движений, которые человек совершает при использовании устройства. Самая простая и распространённая операция, которую уже поддерживают многие гаджеты - это автоматическое изменение ориентации экрана при повороте устройства. Если в современном мобильном телефоне есть акселерометр - скорее всего, он встроен для этого.

Проектировщики и производители карманных устройств шаг за шагом расширяют "язык жестов". Возникает законный вопрос, как необходимо проектировать управление жестов, чтобы пользователи не страдали от такой системы. Какие ключевые особенности должны быть в первую

очередь проработаны, каким должен быть набор жестов управления, как спроектировать управление на базе жестов. Ответам на эти и другие вопросы посвящена данная работа.

В работе рассматриваются вопросы связанные с определением компьютерных жестов, классификации жестовых систем по типу ввода (с описанием особенностей), а также домены применения жестов. Этот материал будет полезен при разработки жестов для конкретного устройства. Рассматриваются вопросы взаимосвязи невербальных жестов и жестов в человеко-компьютерном взаимодействии, связь жестов и контекста (например, связь между жестами и речью). Приведенные классификации и терминологии жестов наиболее часто встречаются в литературе по жестам в человеко-компьютерном взаимодействии и более менее устоявшимися на период написания работы. Большое внимание уделяется характеристикам систем, под управлением жестов. В результате анализа литературы из области юзабилити, эргономики, дизайна и когнитивной психологии были сформулированные базовые принципы критичные, по мнению автора, в жестовых системах. Каждый принцип описывается с точки зрения разработки жестовых систем, снабжен примерами и вариантом реализации. Кроме это, в работе представлена специальная глава, посвященная проектированию жестов в рамках человеко-ориентированного процесса проектирования.

ОБЗОР СИСТЕМ ПОД УПРАВЛЕНИЕМ ЖЕСТОВ

В данной главе приводится обзор и анализ систем под управлением жестов по способу ввода, а также показано развитие пользовательского интерфейса в сторону более естественного способа взаимодействия по средствам жестов.

Способ ввода

Можно говорить о контактном способе ввода (сенсорный ввод, сенсорные датчики, акселерометр, компьютерная мышь) и о бесконтактном способе ввода (компьютерное зрение, remote sensors).

Сенсорный ввод

В общем смысле, под сенсорным вводом понимают средство позволяющее реагировать на прикосновение (пальцем, стилусом и т.п.). Наиболее популярными устройствами сенсорного ввода являются сенсорные панели и/или сенсорные экраны.

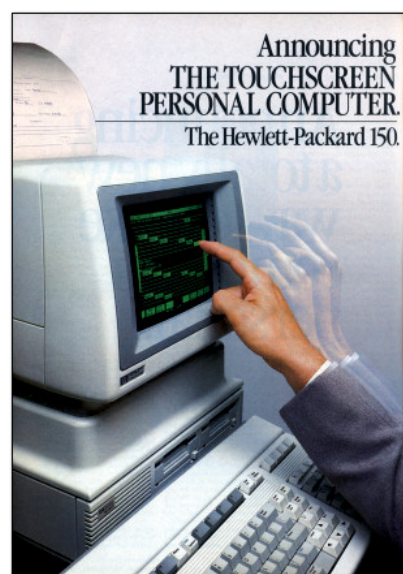
За 40-летнюю историю развития сенсорных экранов было разработано несколько типов этих устройств ввода, основанных на различных физических принципах, которые используются для определения места касания. В настоящее время наибольшее распространение получили два типа дисплеев — резистивные и емкостные. Помимо этого различают экраны, способные регистрировать одновременно несколько нажатий или только одно.

В 1971 году доктор Сэмюел Херст из Университета штата Кентукки изобрел «Элогограф» (Elograph) — сенсорную панель, которая хоть и не была похожа на современные прозрачные экраны, но стала первой важной вехой в истории развития всей технологии. Настоящий сенсорный экран доктор Херст представил миру уже через три года, и его разработка тотчас вошла в список ста важнейших продуктов 1974 года. А еще через три года, в 1977-м, его компания Elographics запатентовала пятипроводный резистивный экран, являющийся одним из самых популярных на сегодняшний день. В феврале 1994 года фирма Elographics сменила название на Elo Touch, под которым известна и по сей день. Более подробно про изобретение и развитие Elograph можно прочитать на сайте компании Elo TouchSystems в разделе История (<http://www.elotouch.com/AboutElo/History/default.asp>).



Первым компьютером с сенсорным экраном принято считать HP-150. Он появился в 1983 году и использовал рамку с сеткой инфракрасных лучей, надетую на девятидюймовый монитор. Эта технология, хоть и не является «по-настоящему» сенсорной, активно используется до сих пор. Главным образом благодаря неприхотливости и ремонтпригодности.

Экран с поддержкой мультитача появился в 1984 году. Это была емкостная прозрачная панель, наложенная поверх ЭЛТ-монитора. В ее создании принимал участие Боб Бойе (Bob Boie) из Bell Labs, хотя саму технологию изобрел Билл Бакстон (Bill Buxton) из Университета Торонто. В одном из интервью Билл отмечал, что «причина, по которой мы сами не сделали сенсорный экран, была не в том, что мы не хотели, мы просто не знали, как сделать его достаточно дешевым». В этом деле и помогла Bell Labs.



Обзор touch технологий

Ниже представлен краткий исторический обзор touch технологий.

- 1954
 - Хью Ле Кейн (Hugh Le Caine) демонстрирует прототип сенсорного органа (Touch Sensitive Organ)
- 1972
 - Сенсорный экран используется в системе PLATO в Университете Иллинойса
- 1974
 - Сэмюел Херст разрабатывает первый прозрачный сенсорный экран
- 1977
 - Elographic запатентовала технологию пятипроводного резистивного сенсорного экрана
- 1979
 - 3M запатентовала технологию поверхностно-емкостного сенсорного экрана
- 1981
 - Тактильные датчики для робототехники
- 1982
 - Elo Touch Systems запатентовала резистивный сенсорный экран
 - Flexible Machine Interface - первый интерфейс, поддерживающий мультитач
- 1984
 - Bell Labs представили сенсорный экран поддерживающий более 1 точки касания - мультитач
- 1985
 - Zenith запатентовал сенсорный экран на поверхностных акустических волнах (ПАВ)

- Input Research Group в Университете Торонто разработали multi-touch tablet
- 1987
 - Elo Touch Systems приобрела ПАВ технологию Zenith
- 1990
 - Nissha разработали сенсорную панель FineTouch
- 1993
 - Apple представил MessagePad H1000 на платформе Newton
- 1996
 - Palm представила Palm Pilot 1000, первый успешный коммерческий карманный компьютер с резисторным сенсорным экраном
- 2001
 - Zytronic разработала проекционно-ёмкостный сенсорный экран и применили его в прототипе планшетного ПК
 - Microsoft демонстрирует прототип планшетного ПК
 - MERL разработала Diamond Touch мультитач систему с использованием технологии проекционно-ёмкостного сенсорного экрана
- 2002
 - Sony опубликовала информацию о мультитач системе SmartSkin
- 2003
 - NextWindow и Smart Technologies начинает производство оптических сенсорных экранов
 - FingerWorks разработала периферийное мультитач устройство ввода
- 2004

- Nintendo производит Nintendo DS карманную игровую консоль с двойным экраном и сенсорной панелью
- 2005
 - Джефф Хан (Jeff Han) демонстрирует FTIR multi touch
- 2007
 - Apple представляет iPhone
 - AUO представили in-cell технологию сенсорной панели на конференции Yokohama FPD
 - Microsoft представляет Microsoft Surface
- 2008
 - HP выпустила планшетовый мультитач
- 2009
 - Dell производить мультитач ПК
 - Основана компания Touchco, разработчик дисплеев по технологии IFSR (Interpolating Force-Sensitive Resistance). IFSR-экраны поддерживают функцию мультитач и способны распознавать несколько степеней давления на поверхность дисплея
- 2010
 - Apple выпустила iPad
 - Amazon приобретает Touchco, производителя гибких дисплеев

Современные технологии сенсорных экранов используются в информационных киосках, банкоматах, различных терминалах, мобильных телефонах и тд. Мы ожидаем встретить чувствительный к нажатиям экран повсеместно, так как это действительно удобно и естественно.

Сенсорные датчики

Использование датчиков было одним из первых методов захвата движение тела и рук для использования в жестовых интерфейсах. Этот метод позволяет отслеживать положение и ориентацию тела человека в пространстве, что значительно расширяет допустимые жесты. Не ограничиваясь плоскостью, мы можем описать непосредственные движения рук, пальцев, головы, корпуса и тд. В 1980 Volt использовал датчики для отслеживания направление руки. Однако, сенсорные датчики довольно дорогостоящие, поэтому их использование не может быть доступным и повсеместным. Подобный метод используется в средах виртуальной реальности.

Более дешевым, является использование электронной перчатки, однако она не сможет передать достаточную точность, которая важна для указаний или манипуляций.

Акселерометр

Наиболее дешевый способ определить положение в пространстве, это использование специального прибора - акселерометр, который позволяет измерять проекцию кажущегося ускорения. Электронные акселерометры часто встраиваются в мобильные устройства и применяются в качестве шагомеров, датчиков для определения положения в пространстве, автоматического поворота дисплея и других целей. В игровых приставках акселерометры используются для управления без использования кнопок — путем поворотов в пространстве, встряхиваний и т. д.

Дешевизна и общедоступность (согласно прогнозу iSuppli, в 2010 году акселерометр будет встроен в каждый третий новый телефон; сегодня датчик движения реализован в одном из пяти аппаратов, в 2008 году он был в каждом одиннадцатом устройстве.), позволяют использовать жесты для решения повседневных задач. Подобные устройства не позволят использовать весь спектр жестовых возможностей человека для передачи информации, но если их правильно использовать для решения определенного круга проблем, то интерфейсы на базе устройств с акселерометром могут быть весьма эффективными.

Компьютерная мышь

Компьютерная мышь - одно из старейших указательных устройств ввода, обеспечивающее взаимодействие пользователя с компьютером. Современные браузеры и некоторые программы поддерживают жесты мышью (анг. *mouse gestures*). Идея в том, чтобы заменить команды в меню на ввод команд с помощью знаков, образованных движением мыши. "Рисовать" команды значительно быстрее и проще, чем искать нужный пункт меню.

Жест фиксируется, если нажать и удерживать старт-кнопку (обычно это правая кнопка мыши) и одновременно «начертить» мышью нужную фигуру. В зависимости от программы движение может отображаться на экране в виде следа (StrokeIt) или не отображаться (Opera).

Многие современные программы (не только браузеры) поддерживают жесты мышью, однако они не столь популярны. С одной стороны, причина не популярности в том, что нет единых общепринятых мышинных жестов (различные браузеры используют различные жесты для одного и того же действия); с другой - предлагаемые системы жестов необходимо специально учить и запоминать, чтобы использовать.



Однако жесты мышью имеет и успешное применение - в играх, например в Black & White. Black & White — стратегическая компьютерная игра, один из наиболее характерных представителей жанра симулятор бога. Разработана компанией Lionhead Studios в 2001. Интересной особенностью игры является возможность сотворения чудес. Чудеса в Black & White — аналог магии, присутствующей во многих играх. Вызов чуда требует начертания в воздухе либо на земле определённой фигуры.

У этой игры появилась целая армия поклонников, но для многих выполнить чудеса (рисование жестов) было довольно сложно. Но так как это игра, то это было еще одной своеобразной “прокачкой skill”. Отсюда можно сделать некий шуточный вывод: если разрабатывается система just4fun, то использование жестового взаимодействия будет плюсом, даже если будет тяжелым для выполнения. Позднее будет создана специальная перчатка для управления виртуальной рукой.



В настоящий момент существуют несколько библиотек и алгоритмов для распознавания жестов мышью, в том числе и на JavaScripts, Flex, что делает возможным использовать жесты мышью на веб-сайтах.

Компьютерное зрение

Компьютерное зрение активно используется в поддержке человеко-компьютерного взаимодействия, в частности в распознавание естественных жестов. Возможно это единственный способ захвата и распознавания системы естественных жестов человека.

В середине 1970 Myron Krueger создал Videoplace — искусственная реальность, окружающая пользователей, которая реагировала на их движения и действия. Это работа ляжет в основу его цитируемой книги Artificial Reality. Videoplace не требовала очков или специальных перчаток, использовались только проектор и камера. Перемещение пользователей отображались цветными силуэтами на экране. Экранные силуэты не были четкими, но давали пользователям чувство присутствия при взаимодействии с другими объектами на экране. Этого было достаточно, чтобы пользователь отстранился, когда его силуэт пересекается с другими силуэтами.

Успехом Videoplace было распознавание естественных жестов, пользователям не нужна была какая либо специальная подготовка, чтобы использовать систему. Как будет рассмотрено дальше, эта система

относиться к типу естественных жестовых интерфейсов (англ. *Natural Gesture Interfaces, NUI*). Рисунки показывают примеры жестов, которые были в системе. На первом рисунке показан своего рода "*Finger Painting*": происходит распознавание указательного пальца, и когда это сделано он становится инструментом для рисунка. Сжатая в кулак рука, позволяет свободно перемещаться, не рисуя на экране. На следующем рисунке показан выбор из меню (в данном случае букв алфавита). Но это можно использовать и как систему ввода текста.



Последняя иллюстрация показывает манипуляцию с объектом (формой объекта), используя указательный и большой палец обеих рук, это управление кривой Безье. Собственно, это современный компьютерный жест *Zoom In/Out*, который очень популярен в обработке изображений на сенсорных мобильных телефонах.



Одним из интересных (и важных) вопросов в системах непрерывных жестах, является *сегментация*. Если посмотреть на рисунок выше, то нам необходимо знать, когда установить форму кривой. Обычное решение, это использование временных задержек — если удерживать руки в одном состоянии некоторое время, издается звуковой сигнал, что форма была изменена. Пользователь может теперь делать что-нибудь другое. Это взаимодействие похоже на то, когда мы держим что-то на месте, пока ждем, когда высохнет клей.

Remote sensors

В 1919 году русский изобретатель Лев Сергеевич Термен представил новый музыкальный инструмент. Звук на этом инструменте возникает не от касания, а только от движений рук исполнителя в пространстве перед

специальными антеннами. При этом со стороны кажется, что звук возникает из ниоткуда. Игра на инструменте заключается в изменении музыкантом расстояния от его рук до антенн инструмента, за счёт чего изменяется ёмкость колебательного контура и, как следствие, частота звука. Вертикальная прямая антенна отвечает за тон звука, горизонтальная подковообразная — за его громкость. Музыкант мог исполнить любое музыкальное произведение, а также создать различных звуковые эффекты (пение птиц, свист и др.). Инструмент был назван по имени изобретателя - Терменвокс. Терменвокс успешен, так как наличие непрерывной обратной связи с движением руки позволяет музыканту быстро построить ментальную модель, как пользоваться устройством. Это хороший пример способности человека, точно управлять системой естественными жестами.



Использование технологий дистанционных датчиков, или *remote sensors* популярно при создании интерфейсов для захвата жестов и движений. В 2007 году вышла игровая консоль Wii от Nintendo, в которой использовался уникальный стиль взаимодействия пользователя и игровой консолью: пользователь физически связывался с игрой. Nintendo утверждает, что функции Wii:

"Several advanced, revolutionary features. Wireless motion-sensitive remote controllers, built-in Wi-Fi capability, and a host of other features have made the Wii the best-selling latest generation console system in the world."

Nintendo (2009)

В 2009 Nintendo продала более 59,16 миллиона консолей. Wii является самой популярной консолью в мире. Причиной такого успеха является интуитивно понятные жесты взаимодействия между пользователем и

игрой, которые реализуются с помощью уникального контроллера — "Wii Remote", который может определять своё перемещение и ориентацию в трёхмерном пространстве. Кроме того, в контроллер встроен динамик и вибро-механизм, что даёт дополнительную обратную связь. Питается контроллер от двух АА батареек. Находясь в режиме ожидания (оранжевый сигнал на кнопке Power), консоль потребляет минимум энергии, но в то же время может получать обновления и сообщения через Интернет, подключаясь к WiiConnect24, сервису, созданному Нинтендо для распространения обновлений, коммуникации владельцев консоли (письма, фотографии) и возможности играть в режиме онлайн. Дискковод консоли начинает светиться голубым светом каждый раз, когда приходит новое сообщение или консоль закачивает обновления.

Главной особенностью Wii Remote по сравнению с конкурентами является детектор движения, позволяющий игрокам управлять персонажами игры и предметами на экране движением руки, или "указывая" на объекты, отображаемые на экране. Это обеспечивается работой акселерометра и светочувствительной матрицы.



Кроме этого, к контроллеру можно подключать различные устройства, расширяя таким образом его возможности. Wii Remote был впервые продемонстрирован 16 сентября 2005 года на выставке Токуо Game Show. С тех пор к контроллеру было привлечено много внимания из-за его уникальных особенностей и большого отличия от конкурентов. Wii Remote обладает способностью восприниматься в трёх измерениях с использованием ADXL330 акселерометра. Wii Remote оснащён оптическим датчиком PixArt, который позволяет определить куда указывает контроллер. Подключение к игровой консоли беспроводное, по протоколу Bluetooth.

Специальная панель (Wii Sensor Bar) является источником инфракрасного света, который воспринимается Wii Remote. Она устанавливается выше или ниже экрана монитора/телевизора (или другом удобном месте, главное, чтобы сигнал попадал в зону действия). Сам

контроллер может комбинироваться со многими выпущенными специально для Wii аксессуарами (рули, пистолеты, ракетки, кии, удочки) чтобы сделать процесс игры более интересным.

Пользовательский интерфейс

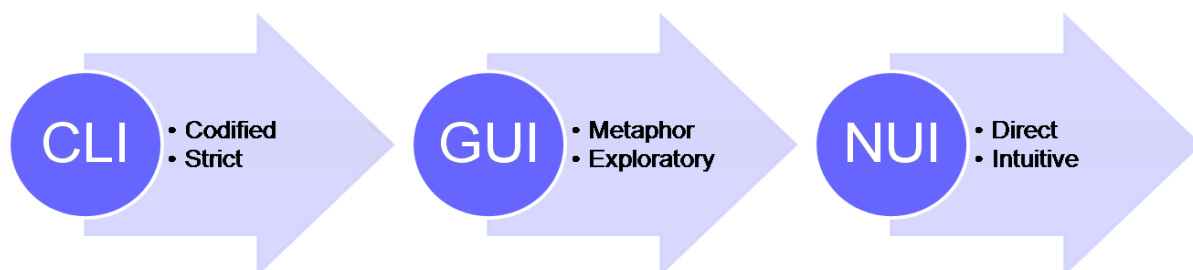
...в интерфейсах не должно быть глаголов...

Под интерфейсам пользователя понимают совокупность средств и методов, при помощи которых пользователь взаимодействует с различными сложными устройствами: компьютер, телефон, автомобиль, терминал и так далее.

В 1968 году, когда Стэнли Кубрик снимал "Космическую одиссею 2001 года", самым распространённым способом взаимодействия с компьютером был диалог: человек давал команду или делал запрос, а машина отвечала. Именно так происходило общение с компьютером HAL-9000 из фильма, только клавиатуру с дисплеем и АЦПУ заменял голос. Реальный 2001 год оказался совсем иным. Синтезировать и распознавать речь к тому времени научились, но для общения с компьютером эти технологии негодились. Популярные пользовательские интерфейсы строятся уже не на диалоге с машиной, а на манипуляциях виртуальными объектами в очень условной двумерной среде. Но и представления о будущих интерфейсах тоже изменились, теперь это не просто более новый старый GUI, это что-то другое.

В последнее время все больше специалистов в области человеко-компьютерного взаимодействия и смежных областях, говорят о сдвиге парадигмы в сторону *более* естественных интерфейсах. На рисунке приведена эволюция пользовательского интерфейса от интерфейса командной строки (англ. *Command line interface, CLI*) до естественного пользовательского интерфейса (англ. *Natural user interface, NUI*). Естественный пользовательский интерфейс характеризуется прозрачностью, невидимостью взаимодействия со стороны пользователя. В интерфейсах не должно быть глаголов, все действия должны быть

понятны и естественны: поднести к уху телефон, чтобы принять звонок, а не нажимать на специальную кнопку.



"Естественное соответствие, то есть использование естественных аналогий и культурных стандартов, ведёт к мгновенному пониманию", - утверждает Дональд Норман. Однако, если человек знает, для чего нужна настоящая мусорная корзина, это вовсе не значит, что он так же легко поймёт, как использовать корзину Windows.

"GUI не способен охватить всё богатство человеческих чувств и умений, которые люди выработали при взаимодействии с материальным миром", - писал в 1997 году профессор Хироси Исии из медиа-лаборатории Массачусетского технологического института (MIT). Он считает, что опосредованные манипуляции виртуальными объектами, из которых состоит графический интерфейс - это тупиковый путь. Будущее за естественным взаимодействием с материальными предметами, надо только найти способ вернуться к нему. Компьютерные интерфейсы, основанные на таком принципе, Исии назвал *материальными*.

Самые интересные особенности интерфейса iPhone тоже основаны на принципе непосредственного взаимодействия с объектами. "Люди не понимают, что мы изобрели новый класс интерфейсов", - говорил Джобс, а затем пояснял: всё дело в том, что в интерфейсе iPhone почти отсутствуют глаголы. Особенность, к которой пытался привлечь внимание глава Apple, действительно очень важна. Пользователь традиционного графического

интерфейса сначала выбирает объект, а затем указывает в меню действие, которое над ним необходимо совершить - тот самый "глагол", о котором говорил Джобс. Пользователю материального интерфейса меню с "глаголами" ни к чему: он берёт объект и сам, своими руками делает с ним то, что нужно. Так работают мультитач-жесты, воспринимаемые эппловским телефоном, так работает прокрутка в браузере и картах iPhone, так работают инерционные списки.

Материальный пользовательский интерфейс

Материальный интерфейс пользователя (англ. *Tangible User Interface*) — это разновидность интерфейса пользователя, в котором взаимодействие с цифровой информацией происходит с помощью материальной конструкции.

Профессор Хироси Исии из Массачусетского Технологического Института — один из первопроходцев материальных интерфейсов пользователя и глава группы разработчиков материальных систем. Его особое видение материальных ИП, т. н. материальные биты — это попытка придать цифровой информации физическое обличие, делая биты осязаемыми и потому доступными напрямую. При этом преследуется цель неразрывно связать такие разные сущности, как миры битов и атомов.

Примером материального ИП можно назвать шаровой автоответчик Дюрелла Бишопа. Каждый шарик соответствует сообщению, оставленному на автоответчике. Перемещение шарика в специальную выемку воспроизводит связанное с ним сообщение или вызывает звонившего.

Другой пример — система Топобо. Блоки её напоминают элементы конструктора LEGO, которые могут соединяться вместе, но, в то же время, перемещаться самостоятельно за счёт моторчиков. Можно тянуть, толкать или поворачивать эти элементы; они запомнят эти действия и способны их воспроизвести.

Другая реализация позволяет пользователю материальным пером набросать рисунок на столе системы. Используя предварительно запрограммированные «жесты», можно клонировать картинку или

растягивать по двум осям, как в программе рисования.

Материальный интерфейс характеризуется:

- Физические представления связаны с основной цифровой информацией;
- Физические представления включают механизмы для интерактивного управления;
- Физические представления считаются и соответствуют цифровым представлениям.

Элементы материального интерфейса - не кнопки и окна, нарисованные на экране, а материальные предметы. Например, целиком телефон. Уже сейчас появляются устройства, которые интерпретируют определенные движения как команды. Некоторые модели Sony Ericsson перемешивают музыку в плеере, если встряхнуть телефон, или прерывают сигнал будильника, если *помахать над гаджетом рукой*. Samsung i900 или Nokia 5800 отключают звонок, когда мобильник *перевернут лицевой стороной вниз*. Разумеется, материальные интерфейсы годятся далеко не всегда, но многое с ними становится значительно проще.

metaDESK

Горизонтальная поверхность metaDESK представляла собой большой проекционный дисплей, над которым на подвижном кронштейне с несколькими степенями свободы перемещался дополнительный жидкокристаллический монитор ("активная линза"). Параллели между традиционным графическим интерфейсом и материальным интерфейсом metaDESK довольно прямолинейны. Дисплей на кронштейне - аналог окна, миниатюрные фигурки, которые расставляются по столу - физические иконки, элементы управления - специализированные инструменты. Конструкция напичкана оптическими, механическими и электромагнитными датчиками, которые отслеживают положение в пространстве каждой части системы.

Тестовое картографическое приложение для metaDESK выводило на поверхность стола спутниковый снимок окрестностей MIT. Пара игрушечных домиков, поставленных на metaDESK, привязывались к реальным строениям на карте.



Перемещая и поворачивая их, можно масштабировать, вращать и двигать изображение. ЖК-дисплей, подвешенный над столом на кронштейне, выводит трёхмерное представление тех же данных. Чтобы увидеть объёмную модель ландшафта и строений в определённой точке карты, нужно передвинуть дисплей в соответствующее место и посмотреть "сквозь" него на стол, как через линзу.

Использование жестов в интерфейсе

Основные элементарные операции, на которых строится GUI - клик, двойной клик, перетаскивание, язык же материального интерфейса более богатый, он состоит из движений, которые человек совершает при использовании устройства. Самая простая и распространённая операция, которую уже поддерживают многие гаджеты - это автоматическое изменение ориентации экрана при повороте устройства. Если в современном мобильном телефоне есть акселерометр - скорее всего, он встроен для этого. Однако ему находят и другие применения. Производители карманных устройств вошли во вкус и шаг за шагом расширяют "язык жестов", который понимает их продукция.

Мобильники Sony Ericsson перемешивали музыку при встряхивании задолго до появления iPod nano со встроенным акселерометром. Другой жест, постепенно получающий распространение, служит для отключения звука. Чтобы заставить замолчать звонок Samsung i900 или Nokia 8800 Arte, достаточно перевернуть устройство лицевой стороной вниз. В некоторых "раскладушках" Sony Ericsson с той же целью изобретательно применена

встроенная камера: телефон затихает, если помахать над ним рукой. В Samsung запатентовали сложную систему знаков, которые устройства смогут распознавать при помощи всё той же камеры.

Более разумным кажется подход компании Gesturetek, которая сооружает интерактивные дисплеи с середины восьмидесятых. Несколько лет назад она заинтересовалась мобильными телефонами и разработала программный "движок" под названием EyeMobile. С его помощью отслеживать перемещения устройства в пространстве можно даже в том случае, когда акселерометра нет. Информация о направлении и скорости движения извлекается из видеосигнала, получаемого со встроенной камеры.

В Gesturetek не стали изобретать собственный язык знаков и попытались придать дополнительный смысл естественным жестам, которые владельцы телефонов делают и так. Например, характерное вертикальное движение, которым мобильник подносят к уху, программа интерпретирует как команду снять трубку, а чтобы перелистнуть страницу, устройством нужно тряхнуть в интересующем направлении. Эту технологию используют в нескольких телефонах, которые продаются в Японии (правда, главным образом, для мобильных игр), но у Gesturetek уже есть варианты этого движка для Windows Mobile и Symbian и даже приложения, поддерживающие управление с помощью EyeMobile: браузер, карты и галерея.

ОБЗОР ЖЕСТОВ ВЗАИМОДЕЙСТВИЯ

В данной главе рассматриваются различные классификации жестов, некоторые из которых используются в психологии невербального общения, а другие в человеко-компьютерном взаимодействии.

В части посвященной жестам в человеко-компьютерном взаимодействии приводится общепризнанное современное понимание и терминология жестов.

Таксономия жестов человека

Жест – знаковое движение рук, ног или головы. Возможно широкое и узкое толкование термина «жест». В широком смысле это слово объединяет все движения человеческого тела (даже мельчайшие и незаметные глазу), а также позы и даже расстояния между собеседниками; в узком смысле жестами называют только движения рук. Использование тех или иных жестов зависит от культуры, контекста и т.д., но они всегда присутствуют во время общения. Это подтверждается тем, что люди обычно жестикулируют во время телефонного разговора, хотя и не видят собеседника.

Когда я работала над этой главой, я хотела найти некую общую классификацию жестов, однако мне не удалось это сделать. Причиной стало то, что в психологии невербального общения еще не выработана общепризнанная классификация жестов. Известные классификации построены на различных основаниях, но даже собранные вместе, они не позволяют во всей полноте обозначить связи жестов со структурой личности и ее общением. Согласно самой широкой классификации, жесты подразделяются на естественные (спонтанные) и искусственные (жесты глухонемых, профессиональные жесты дирижеров, биржевиков и др.).

Однако есть и более узкие, специальные классификации. Я представлю несколько полезных классификаций для будущих рассуждений.

На самом деле многие исследователи не классифицируют, а просто описывают жесты и возможные значения этих жестов. Так, исследователь невербального поведения *А. Штангль* выделяют:

- *жесты рук*: руки, вяло свисающие вдоль тела, — пассивность, недостаток воли; скрещенные на груди руки — тенденция к дистанцированию, известная изоляция, выжидание и др.;
- *жесты кистей рук*: открытая ладонь обращена вверх — объяснение, убеждение, открытое представление, отдавание; руки спрятаны в карманах — скрывание неуверенности, потеря непосредственности; рука сжимается в кулак — концентрация, стремление к самоутверждению и др.;
- *"игры рук"*: пальцы барабнят по столу — демонстрация незаинтересованности; стирающие движения по лбу — стирание нехороших мыслей, плохих представлений и др.; жесты пальцев: палец засунут в рот — инфантильная наивность, удивление, рассеянность; палец плотно прижат ко рту — стремление предотвратить всякое проявление, импульс к овладению собой; палец касается глаз или ушей — опосредованный знак неловкости, плохое настроение, известная робость и др.;
- *рукопожатие*: твердое, энергичное, тяжелое, давящее, бесцветное, слабое и др.

Австралийский исследователь *А. Пиза* представлены наиболее распространенные, с его точки зрения, жесты и действия:

- *жесты ладони*: открытая ладонь, говорящая о правдивости, честности, преданности; пальцы собраны в кулак и виден только указательный палец — «указующий перст», сила, принуждение к действию и др.;
- *жесты кистями и руками*: рука, охватывающая запястье, говорит о неоправданных ожиданиях и попытке самоконтроля; скрещенные руки с большими пальцами, направленными вверх — оборонительное или негативное отношение к чему-либо и др.;
- *жесты "рука к лицу"*: "рука к носу" указывает о недоверии к говорящему; потирание глаза — попытка блокировать ложь;

почесывание шеи — сомнение, неуверенность и др.; рука поддерживает голову — скука, отсутствие интереса; поглаживание подбородка — принятие решения; потирание тыльной стороны шеи — сигнал неоправданных ожиданий и др.;

- *барьеры из рук*: скрещенные на груди руки с ладонями, сжатыми в кулаки, говорят о враждебном отношении к партнеру; дотрагивание до предмета, расположенного вблизи другой руки, — стремление к достижению безопасности и др.;
- *барьеры из ног*: перекрещенные руки и ноги указывают на негативное отношение, попытку защититься от чего-то и др.

Более полезными, на мой взгляд, являются функциональные классификации жестов. Так можно рассмотреть классификацию разработанную психологом *Е.А. Петровой*, в которой учтены не только особенности жестов, но и разнообразные функции *общения*, которые они реализуют:

- *аффективно-коммуникативная* — жесты, выражающие чувства, волю, желание, другие состояния (эмотивная функция); жесты, выражающие течение перцептивных, мимических, интеллектуальных процессов (функция выражения процессов); жесты, сигнализирующие об отношениях, установках, оценках, самооценках (модальная функция);
- *регулятивно-коммуникативная* — фатические жесты (жесты вступления в контакт); конативные жесты (жесты, способствующие поддержанию и усилению контакта); эндные жесты (жесты завершения контакта);
- *информативно-коммуникативная* — презентация информации об объекте, о себе, о другом.

Соотношение вербальной и невербальной информации в процессе коммуникации представлено в следующей классификации *Н.И. Смирновой*:

- *коммуникативные жесты*, замещающие в речи элементы языка (жесты приветствия и прощания, угрозы, привлечения внимания, подзывающие, приглашающие, дразнящие, жесты утвердительные,

отрицательные, вопросительные, выражающие благодарность, примирение и др.);

- *описательно-изобразительные жесты*, сопровождающие речь и теряющие смысл вне речевого контекста (жесты, обозначающие размер, форму предмета, пространственное расположение объекта и др.);
- *модальные жесты*, выражающие оценку предметов, явлений, людей (жесты одобрения, неудовольствия, недоверия, неуверенности, растерянности, отвращения, радости, восторга, удивления).

Чтобы подчеркнуть разнообразие информации, предоставляемой жестами партнеру, их подразделяют на следующие *группы*:

- *жесты-регуляторы*, выражающие отношение говорящего к чему-либо или к кому-либо (кивок, целенаправленные движения руками и др.);
- *жесты-эмблемы* — своеобразные заменители слов или фраз в общении (поднятый вверх большой палец и др.);
- *жесты-иллюстраторы* — жесты сообщения, образные картины сообщения («вот какого размера», «вот такой формы», движения руками, соединяющие воображаемые предметы и др.);
- *жесты-адаптеры*, демонстрирующие специфические привычки человека, связанные с движениями рук (почесывание, поглаживание, перебирание отдельных предметов и др.);
- *жесты-афффекторы*, выражающие через движения тела определенные эмоции (растерянность, удивление, отвращение, радость);
- *жесты оценки* получаемой информации (один палец отставлен, остальные под подбородком — критическая оценка; почесывание подбородка — обдумывание предстоящего высказывания; почесывание пальцем спинки носа — озабоченность, сомнение).

Жестам всегда присуща, во-первых, *вариативность* (даже в тех случаях, когда они общеупотребительные), обусловленная прежде всего индивидуальными особенностями человека; во-вторых, *многофункциональность*. Например, взмах руки может употребляться как

знак отчаяния, привлечения внимания или отказа от чего-либо. Кивки головой не всегда означают согласие — часто они лишь показывают говорящему, что его слушают и готовы слушать дальше, и как бы дают разрешение собеседнику продолжать речь, поэтому в групповой беседе говорящий обычно обращает свои слова непосредственно к тем, кто постоянно кивает.

Вариативность и многофункциональность жестов связаны также с различиями культур, в которых они функционируют: один и тот же жест может быть интерпретирован по-разному. Так, удар по ладони партнера в момент или после произнесения фразы "говорит" египтянину или сирийцу о том, что собеседнику понравилась сказанная шутка, острая фраза, а европеец то же движение рук воспримет как проявление неуважения. Кроме того, жесты, распространенные в одной культуре, могут вовсе отсутствовать в другой. Например, жест "большой палец правой руки опущен вниз", обозначающий у англичан неодобрение, отсутствует у русских, а жест "почесывание затылка", передающий у русских старание припомнить что-то, у англичан встречается довольно редко.

Показатели интенсивности жестикуляции (сила и частота) также определяются культурными нормами. Согласно усредненным данным, на протяжении часового разговора финн прибегает к жестикуляции один раз, итальянец — 80, француз — 120, а мексиканец — 180 раз.

У представителей одной культуры также наблюдаются различные показатели интенсивности жестикуляции вследствие индивидуальных различий. Чрезмерное жестикулирование отмечается у человека, например, при ослаблении или отсутствии обратной связи со стороны партнера по общению; может быть также выражением беспокойства, неуверенности человека; в ситуации переговоров, сопровождающихся сильными эмоциями, более интенсивно жестикулируют те, кто претендует на лидерство.

Таким образом, жесты свидетельствуют об интенсивности переживаний, о качестве и направленности отношений, о культурной и групповой принадлежности. Жесты, как и мимика, могут нести самостоятельную информацию о человеке независимо от его речи.

Также, по мнению Т.М. Николаевой, при классификации жестов необходимо учитывать тот факт, что у человека отсутствует возможность запечатления движения в трехмерном пространстве и времени. В процессе коммуникации движение может быть запечатлено только в одной из плоскостей, поэтому для описания жеста следует указать либо орган, выполняющий движение; либо его направленность; либо цикличность, этапность в совершении движения; либо его отношение к положению человеческого тела в пространстве. Эти принципы описания положены в основу словарей жестов. Вопросы нотации и документации жестов будут рассмотрены в последующих главах.

Классификация жестов человеко-компьютерного взаимодействия

Когда мы говорим о жестах в общении со сложными системами, в том числе и с компьютером, нам интересно рассматривать естественные жесты человека в рамках функций, которые они выполняют. Ниже я буду приводить классификации и терминологии более менее устоявшиеся в смысле жестов в человеко-компьютерном взаимодействии.

В 1994 году Cadoz предложил сгруппировать жесты в три группы по функциям:

- *semiotic*, чтобы сообщить информацию;
- *ergotic*, чтобы управлять физическими объектами;
- *epistemic*, чтобы изучать окружающий мир: тактильное исследование.

Чтобы реализовать поддержку жестов в рамках управления компьютером, необходимо особенно уделить внимание первой группе *semiotic* жестам.

Классификация жестов по функциональности

Идея о более подробной классификации семиотические жестов была рассмотрена в нескольких работах, но наиболее полной, вероятно, это исследование Axel Mulder (1996), касательно использования жестов в

человеко-компьютерном взаимодействии. В своей работе он собрал результаты многих исследователей.

Семиотические жесты часто классифицируют по функциональности, мы рассмотрим таксономию жестов, предложенную Rime и Schiaratura (1991):

- *символические жесты* (англ. *symbolic*) - это жесты, которые в пределах каждой культуры имеют однозначное значение. Хорошим примером может служить жест “OK”, или язык глухонемых, например амслен (American Sign Language, ASL);
- *дейктические (указательные) жесты* (англ. *deictic*) - это жесты указания, или, другими словами, служат для привлечения внимания слушателя к определенным событиям или объектам окружающей среды. Когда говорят о жестах в человеко-компьютерном взаимодействии, в первую очередь вспоминают как раз о дейктических жестах и это не удивительно, эти жесты используются практически во всех доменах человеко-компьютерного взаимодействия (виртуальная реальность, мобильные телефоны и тд). В работе Bolt “Put that there” (1980) дейктические жесты использовались совместно с командами ввода голосом для указания расположения предметов на большом экране;
- *изобразительные жесты* (англ. *iconic*) - эти жесты используют, чтобы передать собеседнику информацию о размере, форме или ориентации объекта беседы. Эти жесты, обычно сопровождаются фразами, например: “Я поймал вот такую большую рыбу”.
- *пантомимические жесты* (англ. *pantomimic*) - это типичные жесты, для показа движения некоего невидимого инструмента или объекта в руке говорящего. Например, когда изображают, как играет ударник.

Put-That-There

В одной из первых работ по мультимодальным системам (R. Bolt “Put-That-There”, 1980) демонстрировалась концепция использования речи и ручного манипулирования объектом. Начиная с работы Bolt появились библиотеки, архитектурные

решения, алгоритмы для поддержания новых видов взаимодействия. Также стали появляться реальные приложения. В отличие от первоначальной концепции Bolt, развилась общая идея мультимодальных систем: взаимодействие с объектом не ограничивается только прямым манипулированием. Кроме того, активная эмпирическая работа по созданию таких систем способствовала построению теории НСМІ (human–computer multimodal interaction), что обеспечивает основу для написания руководства по проектированию таких систем.

Таксономию жестов Rime и Schiaratura расширил McNeill (1992), добавив жесты, касающиеся типа коммуникации:

- *жесты акценты* (англ. *beat*) или *жесты ударения* - эти жесты служат, чтобы расставить ударение в речи, указать на значимые слова и тд.. Так называются, потому что рука ходит вверх вниз, как будто отбивая такт речи.
- *связующие жесты* (англ. *cohesive*) - эти жесты служат для связывания временно отделенных, но тематически связанных частей беседы. Это важная группа жестов, хотя с может показать, что это всего лишь вариация изобразительных, пантомимических и/или дейктических жестов.

Например, если мы хотим сделать переход по ссылке жестом, то этот жест (для пользователя) обязательно должен быть связующим, а не каким либо другим.

Классификация жестов по связи с речью

Приведенная классификация помогает определить, функциональную нагрузку каждого жеста, в отдельности, но в нормальном общении так не бывает, для уверенности в интерпретации информации мы часто используем и жесты, и речь вместе.

Семантические жесты могут интерпретироваться без дополнительной информации о контексте, либо контекст должен быть обеспечен другим

последовательным жестом, действием или речью. Таким образом, семантические жесты могут также характеризоваться относительно связи с речью:

- жесты, из которых *“всплывает”* смысл речи: символические, дейктические;
- жесты, которые *изображают* смысл речи: изобразительные, пантомимические;
- жесты, которые *касаются диалогового процесса*: жесты ударения, связующие жесты;

Многие исследователи искали связь между речью и жестами. Kendon предложил расположить в порядке зависимости речь/жест:

- жестикуляция (жесты ударения, связующие жест);
- языкоподобные жесты (изобразительные жесты);
- мимика, пантонима (пантомимические жесты);
- символы (дейктические жесты);
- язык жестов (символические жесты).

Чем ближе группа к последней группе "язык жестов", тем снижается необходимость в сопровождающей речи, жесты становятся более подобны языку. Например, язык жестов (язык глухонемых) учитывает многое из синтаксических и семантических особенностей речи, поэтому не нуждается в речи для интерпретации. Однако изобразительные жесты не могут быть поняты без дополнительной информации.

Предложенная Rime и Schiaratura и McNeill классификация жестов признанная и используется в большинстве работ по жестам в человеко-компьютерном взаимодействии. Однако, в ней ни как не выделяются домены в которых могут быть использованы жесты.

В работах Quek и связанных исследованиях, выделяют пять основных типов жестов в человеко-компьютерном взаимодействии:

- *жестикуляционные жесты*;
- *манипуляционные жесты*;
- *сигнализирующие жесты*;
- *дейктические жесты*;
- *языкоподобные жесты*.

Стоит обратить внимание на не используемые термины. В большинстве литературы по жестам в человеко-компьютерном взаимодействии используются либо классификации McNeill, либо группировка жестов представленная выше.

Манипуляционные жесты

Существуют много различных форм манипуляционных жестов в человеко-компьютерном взаимодействии. Следуя определению Quek, назначением манипулятивного жеста является управление определенным объектом, используя связь между фактическим движением рук / руки и манипуляции объекта.

При манипуляция может происходить, как используя некое устройство манипуляции, как компьютерная мышь, или стилус, так и используя свободные движения рук, для имитации манипуляций физическим объектом; как в интерфейсах виртуальной реальности, так и в “физических” интерфейсах. Можно рассмотреть некоторые типы манипуляционных жестов.

Начнем с наиболее простых и привычных: манипуляции объектами *двумерными жестами*. Двумерные жесты, означают жесты с 2-мя степенями свободы. В каких-то случаях манипулируют непосредственно курсором, окном или другим объектом на экране. Традиционные методы взаимодействия используют мышь, стилус и другие устройства ввода с 2 степенями свободы для манипулирования в GUI. Современные сенсорные экраны тоже относятся к такому способу взаимодействия. Непосредственное взаимодействие часто ассоциируется с настольными приложениями, и включает в себя такие действия, как перетаскивание, движение и выбор объекта (клик), но они не являются жестами. Это не жесты, до тех пор, пока пользователь не выполнит какой-то жест, который будет интерпретировано как команда. Одной из особенностью манипулятивного жеста является наличие параметров в системе, которые позволяют пользователю перетащить, переместить или изменить цифровой объект [Rubine 1992].

Очень важно понимать разницу между жестом, как *команда* и жестом, как *манипуляция*. Для начала можно обратиться к определениям:

Жест (от лат. *gestus* — движение тела) — некоторое действие или движение человеческого тела или его части, имеющее определённое значение или смысл, то есть являющееся знаком или символом.

Манипуляция (от лат. *manus* - рука) — действие рукой или руками при выполнении какой-либо сложной работы; сложный прием в ручной работе. Умело управляться.

Интуитивно понятна разница, но можно выделять 4 основных различия:

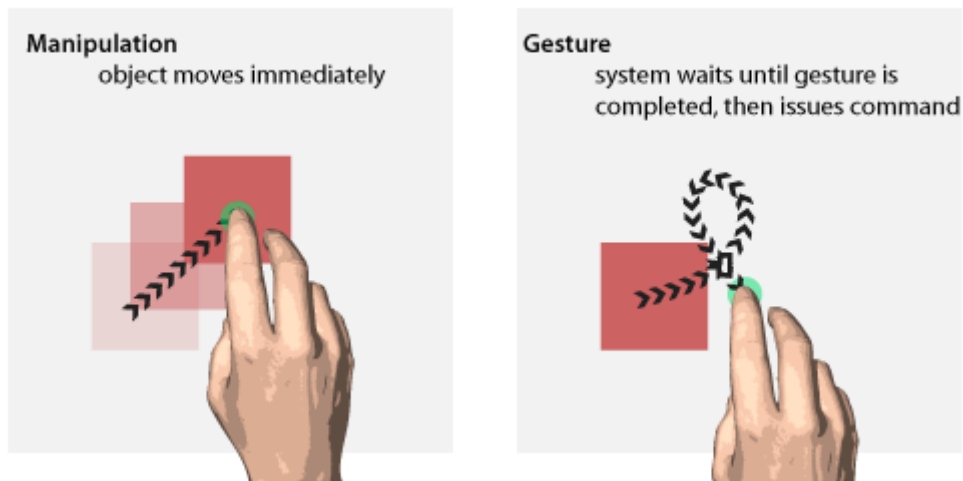
Манипуляция:

- *зависят от контекста*: выполняется только на конкретном объекте/ах и в определенном месте/местах;
- *немедленное реагирование*: есть прямая корреляция в причинно-следственной связи между действиями пользователя и действиями системы;
- *прямая* – действия пользователя непосредственно влияют на работу системы, объект или событие в некотором роде.

Жест:

- *не зависят от контекста*: они могут быть где угодно в системе, не зависят от места или объекта;
- система ожидает *серии событий*, чтобы определить, как реагировать;
- *косвенный* – действия пользователя не влияют на системы непосредственно. Действия в некотором роде символические, вроде команды, выражения.

На картинке ниже схематично показано основное отличие жестов.



В книге Dan Saffer's, *Designing Gestural Interfaces*, (O'Reilly, 2009) она пишет (что считать жестом в книге): *“для целей этой книги, любое физическое движение, которое цифровая система может распознать и реагировать без помощи традиционных устройств ввода, таких как мышь или стилус”*. Это может быть простой способ определить типы взаимодействия для книги, но не является правильным обобщать таким образом. Профессор Шнейдерман в 1983 году дал точное определение *прямой манипуляции*, и был абсолютно прав. Непосредственная манипуляция должна быть только непосредственной (когда мы начинаем обсуждать более сложные движения, которые являются командами, нам нужен новый набор терминов.)

Жесты, как манипуляции являются самыми распространенными и наиболее шаблонными так как их легко проектировать, легко понять, и пользователю интуитивно понятен конечный результат.

Мне хочется добавить еще одно ограничение, касательно различий жестов-команд и жестов-манипуляций. При манипуляции важно, чем ты это делаешь, одной рукой, или двумя; или же используя 3 пальца. При выполнении жеста, не важно, чем ты рисуешь жест (как наподобие клавиатуры: без разницы используется 10-пальцевая слепая печать, или же аккуратно вводится символ за символом одним указательным пальцем).

Все что описано выше, это базовое понимание манипулятивных жестов, однако, они не ограничиваются применимостью только на дисплеях компьютеров, мобильных телефонах и тд. Например, манипуляция типа перебрасывания файлов, графических объектов с одного устройства на другое тоже является по большей части манипуляцией. Манипулировать можно и имитируя некий предмет в руке.

Можно сделать вывод: если есть какие то параметры которые напрямую связывают движение рук (не важно посредством какого либо предмета или нет), то это манипуляция.

Семафорные жесты

В 1801 году французский артиллерийский офицер Депиллон изобрел средство связи корабля с берегом, использующее перемещение трех перекладин, поднятых на мачте. Изобретение получило название семафор, что в переводе с греческого означало “знак несущий”. Помимо семафорной азбуки (каждой букве свой знак), также передавались и команды (“Повторить”, “Передвинетесь влево от меня” и тд). Видимо, это является наиболее известными семафорными жестами. Итак, *семафорной*, называется любая система жестов, использующая стилизованный словарь определяющий динамические или статические жесты рук.

Семафорные жесты могут рассматриваться как коммуникабельные, в качестве универсальных символов общения с компьютером.



Семафорные жесты могут рассматриваться как коммуникабельные, в качестве универсальных символов общения с компьютером.

Семафорные жесты довольно популярны, даже несмотря на то, что передача информации при помощи сигналов или знаков является

крохотной долей всего человеческого общения. Семафорные жесты не являются естественными жестам, они всего лишь обеспечивают некую функциональную полезность. Однако с переходом к более повсеместным компьютерным системам, использование семафорных жестов рассматривается как практический способ обеспечить удаленное управление в “умных” комнатах и интеллектуальных средах, а также в качестве концентрации на более важных задачах. Мы опишем несколько категорий жестов, которые попадают под категорию семафорных жестов.

Статические, динамические жесты.

Семантические жесты могут быть как статическими, так и динамическими, в отличие от манипулятивных жестов, которые в основном динамические. Например, когда мы соединяет большой и указательный палец в жесте “ОК”, это статический жест, тогда как махание рукой является динамическим. Эти типы жестов могут быть выполнены используя руку, пальцы, голову, ноги и другие пассивные или активные (электронные) объекты, как карандаш или компьютерная мышь.

Жестовые ударения.

Жестовые удары, это короткие и простые по форме движения руки, обычно согласованные по времени с интонационным ударением, хотя могут появляться независимо от него. Под семафорными жестами такого типа в человеко-компьютерном взаимодействии обычно подразумевают жесты, для перехода по веб-страницам. В интерфейсы страницы, вставлены специальные маркеры, переходя по которым (в случае мышки или стилуса, кликая по которым) мы попадает на новую страницу, тем самым поддерживая поток восприятия информации, подобно как естественные жестовые ударения помогают в речи.

Жестикуляция

Жестикуляция (некоторых источниках, под термином "*coverbal*") в литературе по человеко-компьютерном взаимодействии рассматривается, как более естественная форма жестов, и чаще всего жестикуляцию рассматривают в многомодальных интерфейсах, где речь и жесты рассматриваются как единый способ коммуникации. Хотя известно много о

том, когда и каким образом употребляются те или иные жесты, совершенно неясно, почему люди жестикулируют.

Согласно первой гипотезе, люди жестикулируют просто потому, что видят жесты других людей и усваивают подобную модель поведения. Исследователи изучили видеозапись разговоров 12 слепых от рождения и 12 зрячих людей в возрасте от 9 до 18 лет. Оказалось, что все 12 слепых во время разговора жестикулировали столь же интенсивно, как и зрячие. Более того, в одних и тех же ситуациях и слепые, и зрячие люди часто применяли похожие жесты. Так, объясняя на словах, что жидкость необходимо перелить в другой контейнер, и те, и другие делали одинаковый жест. Это особенно интересно, поскольку слепые люди не могли видеть и скопировать этот жест; они должны были его придумать.

Вторая гипотеза состоит в том, что жесты передают слушателю дополнительную полезную информацию. Для ее проверки ученые провели ряд опытов, в которых слепого человека просили передать некую информацию экспериментатору. При этом половине испытуемых говорили, что экспериментатор тоже слеп, а вторая половина знала, что он зрячий. Оказалось, что в обеих ситуациях интенсивность жестикуляции оставалась практически идентичной.

Существует третья гипотеза, по которой жесты могут отражать или даже облегчать ход мысли говорящего человека.

На мой взгляд, рассмотрение и/или поддержка жестикуляции, как одной из форм жестов, в человеко-компьютерном взаимодействии необходима (только?) в системах, где важна поддержка человеко-подобного взаимодействия, например, во взаимодействии робота и человека, так как робот должен “прикидываться” человеком. В “чистых” жестовых интерфейсах, жестикуляция вряд ли принесет какой-нибудь дополнительный смысл, поэтому мы должны ее учесть, только как возможный шум, так как человек все равно будет жестикулировать.

Язык жестов

Жесты, как грамматически и лексически полный язык, в человеко-компьютерном взаимодействии редко рассматривается в качестве

взаимодействия с компьютером. Это связано, во-первых, с технической сложностью распознавания такого языка и интерпретации, а во-вторых, доступностью – такой язык пользователю необходимо сначала изучить.

Область применения жестов

Применение жестов возможно в большинстве доменах человеко-компьютерного взаимодействия. Во многом, это обосновано тем, что жестикация естественна для людей, и во многом определяется контекстом общения, действия и тд.

К ним можно отнести, например:

- виртуальная, дополненная и другие типы реальности;
- взаимодействие человек-робот (англ. *Human-Robot Interaction, HRI*), системы типа аватар;
- моделирование 3D;
- работа в умных средах;
- игры;
- мобильные телефоны.

Понятно, что практически все области человеко-компьютерного взаимодействия могут применять жесты, так как это, безусловно, более естественный способ взаимодействия. На мой взгляд, жесты будут более успешны (по сравнению с традиционным способом управления) в системах виртуальной реальности и HRI.

ХАРАКТЕРИСТИКА СИСТЕМ ПОД УПРАВЛЕНИЕМ ЖЕСТОВ

Мы выросли на интерфейсах, построенных под управлением мыши и клавиатуры, но уже знаем, что они могут быть "неюзабилити" и способны превратить любую работу в кошмар, а иногда и в катастрофу. В новых интерфейсах ключевым вопросом является построение управления на основе жестов, именно оно может решить многие текущие проблемы, или же породить еще больше проблем.

Управление на основе жестов уже сейчас достаточно популярно: оно присутствует практически во всех современных мобильных телефонах с сенсорным экраном и/или встроенным акселерометром. Мы повседневно сталкиваемся с интерфейсом под управлением жестов. Возникает законный вопрос, как необходимо проектировать управление жестов, чтобы пользователи не страдали от такой системы. Но чтобы ответить на этот вопрос, нам нужно понять какие ключевые особенности должны быть в первую очередь проработаны.

В этой главе я хочу порассуждать о неких характеристиках "хорошей" системы жестов.

Ясная концептуальная модель

В инженерной психологии понятие *концептуальной модели* определяют, как совокупность представлений оператора о реальном и прогнозируемом состоянии объекта деятельности и системы в целом, о целях и способах реализации своей деятельности. Если в основе лежит ясная концептуальная модель, то можно предсказать результат действий с нею. Если такая модель скрыта (или еще хуже, отсутствует) то пользователь работает как сапер: до первой мины. Пока все хорошо, проблем нет, но если пользователь столкнется с новой ситуацией, помочь ему сможет только понимание, основанное на ясной концептуальной модели.

В современной литературе достаточно много уделяется описанию *значимости* концепции (считают, что это самое главное в проектировании

пользовательских интерфейсов), однако не так много методов *построения* ясной концептуальной модели. Важным аспектом при построении является не только анализ списка объектов системы и доступных действий над этими объектами, а возможность донести до пользователя заложенную концепцию, для понимания системы. При проектировании жестовых систем последнее является наиболее важным. Все таки, как можно передать управление, основанное на жестах (особенно на свободных жестах, где отсутствуют какие-либо ограничители)?

Учитывая тот факт, что концептуальные модели являются разновидностью ментальных моделей - моделей, которые складываются в нашем сознании о себе, других, окружающей среде и повседневных задач на основе опыта, практики и обучения можно предположить, что концептуальная модель управления жестов может быть представлена в каких-то случаях некоторой *невербальной метафорой движения* (метафора действия), разновидностью метафоры интерфейса. В общем понимании эта метафора может строиться на любом типе невербальной коммуникации, например на паралингвистике (звуковых кодах: тембр, тон, интонация) или мимике, но нас интересует больше построение компьютерной метафоры на основе жестикуляции человека.

Обычно, когда говорят о метафорах в контексте проектирования пользовательского интерфейса и проектирования взаимодействия, подразумевают визуальные метафоры, здесь же речь идет о метафоричном способе управления, основанном на жестах, которые мы используем в повседневной жизни. В качестве пояснения рассмотрим знакомый пример. В современных мобильных телефонах с поддержкой мульти тач, практически *de facto* стал жест *pinch/zoom* для изменения размера изображения - увеличить/уменьшить. Если попытаться найти функциональную группу жестов для этого компьютерного жеста, то это иконический жест: в общении мы аналогично показываем форму обсуждаемого предмета. Этот жест построен на метафоре невербального общения, что позволяет пользователям построить концептуальную модель управления размером картинки, основанной на знакомом невербальном движении.

Сложность построения невербальной метафоры движения заключается в выборе набора жестикуляций, который будет *как-то* отображаться в управление системы. По сути, эта проблема сводится к проблеме поиска "хорошей" метафоры, конструирования метафоры и смежным вопросам, ответы на которые не является темой данной работы [подробнее можно узнать в работах В. Л. Авербуха]. При построении невербальной метафоры движения не стоит забывать, об этом нам говорит Владимир Лазаревич, что "...метафоры успешны, когда использованные метафоры уменьшают абстрактность компьютерной модели, в том числе абстрактность интерфейса пользователя с системой".

Также необходимо помнить, что если Вы используете символические жесты, то они могут отличаться (в корне) для различных культур. Так Ричард Никсон попал в одну неприятную ситуацию, когда выступал с речью в Бразилии он поднял руку с соединенным большим и указательным пальцем. Как и для многих американцев, для него это был дружелюбный жест "ОК". Но для жителей Бразилии это было сексуальное оскорбление. Кроме того, для жителей Турции этот жест является смертельным оскорблением. Нужно довольно ответственно подойти в выборе набора жестикуляций; лучше всего, если система основана на дейктрических, иконических и пантомимических жестах, в других ситуациях необходимо быть уверенным в том, что контекст использования данного жеста совпадает как и пользователя, так и у системы (так и у проектировщика).

Таким образом, необходимо донести до пользователя концептуальную модель управления, на основе жестов. В модели нужно передать **начало** и **конец** жеста, **выполнение** жеста, **распознавание** жеста. А также важным моментом, является время выполнения жеста или **скорость выполнения**: несмотря на то, что два жеста по траектории движения могут совпадать, они отличные, если выполнены с разной скоростью. Не один из этих моментов не должен отсутствовать или быть не однозначным. Пользователь должен хорошо понять как работает система и чем яснее мы ему это сможем передать, тем корректнее он сформирует свое мнение о том, как управлять системой с помощью жестов. Самый простой способ позволить пользователю построить правильную ментальную модель, это

создать *ограничители*. Ограничители могут быть физические, смысловые, культурные и логические. Например, для систем с сенсорным экраном, ограничителем начала жеста является нажатие на экран. А вот для управления свободными жестами необходимо создавать логическое ограничение. Таким ограничением может служить специальное визуальное отображение выполняемых действий на экране.

Кроме того, необходимо учесть, что пользователям, то есть людям, свойственно ошибаться. Пользователь может случайно сделать какое-то движение, которое система распознает как начало жеста. В этом случае должно быть возможность сказать системе "нет, нет...я не хотел ничего выполнять". Этот вопрос подробнее рассматривается ниже, в характеристике "Возможность отмены".

Соответствие

В данном случае, соответствие означает не только связь между жестовым управлением и результатом выполнения, но и связь между жестовым управлением, как вынужденным действием управления системы и привычными невербальными движениями. Рассмотрим соответствие на примере переключения слайдов, или, что точнее, перелистывание слайдов. Пользователь должен выделить два соответствия:

- что перелистывать презентацию можно жестикулируя определенным образом;
- что перелистывать можно вперед (Следующий слайд) и назад (Предыдущий слайд).

Оба соответствия в определенной мере случайны. Для того чтобы обеспечить естественное соответствие, мы связываем жестикуляцию с знакомым невербальным движением, которое соответствует команде "Следующий слайд" и, обратное движение, которое соответствует команде "Предыдущий слайд". По функциональности - это должен быть связующий жест, который связывает текущий и следующий слайд. В таксономии жестов по функциональности этому соответствует группа жестовых ударений, следовательно наш жест управления должен *как-то*

отображать в группу невербальных жестовых ударений. Это означает, что компьютерный жест должен быть простым, быстрым, ненавязчивым и после его выполнения, мы должны получить связанную информацию. Если это сенсорный экран, то это некое *перелистывание* вправо или влево; если считать, что у пользователя в руках некий контроллер с акселерометром, то данных жест может выглядеть как наклон устройства влево или вправо - *перекат*. Этот пример достаточно прост и очевиден, но он показывает как компьютерный жест управления может быть изоморфен естественному жестовому ударению. И это является причиной удачного проектирования жестов "Следующий слайд" и "Предыдущий слайд". Если добавить моментальную обратную связь (отклик системы на движение рукой), то соответствие будет явным и наглядным.

Группа невербальных жестов может помочь разработать жесты, если компьютерный жест будет обладать теми же характеристиками, что и реальные жесты в рамках своей функциональности. В большинстве случаев метафоричный жест должен соответствовать своему функциональному назначению - не нужно использовать иконические жесты для привлечения внимания к системе, или для командных действий (вроде "Сохранить").

Обратная связь

Согласно принципу обратной связи, пользователь должен получить полную и точную информации о результате проделанных им действий. Наличие хорошей обратной связи позволяет сократить разрыв оценки - усилий, которые уходят на интерпретацию состояния системы и определения связи между ожиданиями и намерением и действительным результатом. Для системы под управлением жестов это принцип является наиболее важным, так как только (!) по обратной связи пользователь может понять какие действия можно выполнять или какие действия были выполнены или не выполнены. Если в системе под управлением свободных жестов нет возможности организовать тактильную обратную связь, мы должны осуществить поддержку хорошей визуальной и звуковой

связи. В сенсорных экранах визуально отображаются нажатые клавиши, это хорошо, но для пользователя с проблемами со зрением это не поможет, да и звук не всегда можно использовать. В большинстве случаев, мы ограничиваемся только визуальной обратной связью, но хорошо, если удастся создать мультимодальную обратную связь.

Обратная связь должна быть не только после распознавания жеста, но как только система понимает, что сейчас идет какой-то командный жест — она должна об этом сообщить пользователю. Это необходимо, так как распознавание жестов не достигает 100%; поэтому важно, чтоб пользователь понимал нужно ли ему еще раз повторить этот жест. Еще одно важное свойство, когда говорим об интерфейсе, это время отклика. Книжки по юзабилити советуют, что время отклика должно быть менее 100ms.

В простом случае, можно посоветовать использовать специальную иконку, которая сообщит пользователю о том что жест начался и то что он распознан, или не распознан. В более сложном, обратная связь должно показывать сам жест и как идет процесс распознавания.

В своей работе я использовала оба вида предложенной обратной связи.

Affordance

Самым точным переводом *affordance* на русский язык будет, пожалуй, то, *каким образом следует взаимодействовать с данным объектом или использовать его*, или *воспринимаемая доступность*. Это качество продукта или системы. Термин "*affordance*" был предложен Джеймс Гибсон (James J. Gibson), основателем направления в психологии, рассматривающего восприятие как процесс, не предполагающий умозаключений, промежуточных переменных или ассоциаций. Позже идея "*affordance*" была развита в книге Дона Нормана "*Дизайн промышленных товаров*" (англ. *The Design of Everyday Things*). Если продукт, система обладает свойством *affordance* это означает, что пользователь сможет обойтись без какого-либо предварительного обучения и сразу станет использовать продукт. В каком то смысле *affordance* может быть критерием понятности.

От жестового управления требуется, чтобы пользователь в первую очередь догадался о том,

- что системой можно управлять жестами;
- какие жесты поддерживает система.

И это пользователь должен понять из самой системы, а не из какого-то дополнительного руководства или чего-то еще. Понятно, что в общем случае нельзя ответить на вопрос, как добиться *affordance* управления жестов. Но можно посоветовать дополнить обратную связь *предсказанием*, то есть в процессе распознавания жеста показать какие могут быть доступны. Это позволит пользователю с одной стороны представлять конечную цель, а с другой видеть какие еще возможности поддерживает система. Так же это может быть полезно для жестов, траектория которых одинаково начинается.

Возможность отмены

Многие исследования пользовательских интерфейсах, неоднократно замечают, что привычные компьютерные интерфейсы не по-человечески работают с пониманием отмены. Люди обычно сначала что-то делают, а потом пытаются исправить, что сделали, и почти никогда не задаются вопросом "а точно ли я хочу это сделать". Причем они могут отменить как только, что совершенное действие, так и действие сделанное в далеком прошлом.

Если думать о жесте, как о команде, то должна существовать специальная команда, которая бы отменяла любую другую команду и в рамках системы этот жест должен быть одинаковым, некий универсальный управляющий жест.

На мой взгляд, жест "*отмена*" дает огромное преимущество по сравнению с обычными интерфейсами. Можно отметить, что жест-отмена это совсем не то, что отмена в меню интерфейса. Я поясню. Когда мы выбираем какой-то пункт в меню, то мы хотим, чтобы программа сделала это действие: программа должна отменить последнюю операцию или, например, программа должна сохранить документ. Когда мы

жестикулируем, это делаем мы сами. Нет никакого посредника, нет никакого разделения полномочий.

Мне нравится реализация отмены для акселерометра, как встряхивание. Вполне интуитивно понятно, и однозначно.

Адаптивность

Про проектировании жестового интерфейса есть две возможности: создать универсальный язык жестов, либо позволить пользователям самим определить набор жестов. В первом случае довольно трудно определить персональные жесты, поскольку многое может зависеть от культуры, индивидуальных особенностей (быть левшой или правшой) и т.д., все это влияет на наши естественные жесты. Поэтому, жестовый интерфейс должен либо приспосабливаться к пользовательским жестам (наблюдение), либо видоизменять универсальные жесты, либо позволить определить персональный набор жестов. Первая возможность, основанная на наблюдении, предлагает обнаружение оптимальных жестов для данной задачи, но это длительный процесс, поэтому может не подойти для повседневного использования. Оставшихся два подхода предпочтительней для повседневных систем. Для реализации необходимо использовать алгоритмы машинного обучения. При автоматической сегментации возникают проблемы алгоритма распознавания, так как захватываются движения, которые не предназначаются жесту. Эти движения считают шумом, и машинные модели обучения должны учитывать это.

По-моему мнению, жесты должны представлять собой изначально какой-то базисный алфавит жестов, причем некоторые жесты уникальны (например, жест отмены) в рамках всей системы, а некоторые не уникальны, то есть один и тот же жест может использоваться по-разному в зависимости от контекста. Но при этом у системы есть возможность адаптироваться, персонализироваться под конкретного пользователя и условия.

А вот насчет возможности пользователям самим определить набор жестов не все так однозначно. С одной стороны это снимет ответственность

за проектирование хорошего набора жестов управления с проектировщика, а с другой заставит пользователя думать над тем как жестикулировать, чтобы выполнить какое-то действие. Другими словами, пользователь должен знать как работает программа, какие жесты она сможет распознать, поддерживает ли она контекст (а иначе все жесты должны быть уникальны), какие действия нужно повесить на жесты и т.д.. Как только речь идет о том, что пользователь должен что-то знать о системе, а в человеко-ориентированных системах пользователь не должен постоянно помнить о том, как устроена эта "штука" [исключением могут быть случаи, когда речь идет о специализированных программах, или пользователь по должностным инструкциям обязан знать как работает система]. Поэтому я считаю, что позволить пользователю самому выбрать набор жестов можно лишь в том случае, если система сможет, на основе наблюдений, предложить пользователю жест.

Использование системы в целом должно быть простым, она должны просто настраиваться и использоваться. Также, необходимо помнить, чтоб пользователи не обязаны знать, как система работает, поэтому необходимо выявлять неправильное использование и игнорировать его. Если система не правильно распознала жест, нужен простой механизм отмены действия и возможность отметить, что жест распознан не правильно. Для того чтобы определить персональные жесты, необходим интерфейс, который помогает определению набора жестов.

Доступность

Понятие доступности (англ. *accessibility*) 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 различными физическими или техническими ограничениями. А также, системы под управлением жестов не должны (в идеале) накладывать ограничение на среду в которой используются. То есть, система должна быть доступна

в любое время, при любом освещении и при любом пространственном положении. В настоящее время, системы свободных жестов в основном базируются на камере или камерах и накладывают ограничения на среду в которой используются. Кроме того, калибровка камеры может быть довольно сложной задачей для обычного пользователя.

Вывод

Чтобы система жестов была удачная, нужно как можно яснее передать концептуальную модель управления. Чтобы этого добиться, необходимо поддерживать с пользователем непрерывную обратную связь, использовать визуальные ограничители, соблюдать принцип соответствия жестов и функциональную нагрузку действия, а также, управление жестов должно быть видимым для пользователя.

В концептуальной модели жеста для любой системы пользователь должен ясно представлять себе, что компьютерный жест характеризуется такими параметрами, как

- начало и конец жеста;
- выполнение жеста;
- скорость выполнения жеста.

Для того, чтобы понять удовлетворяет ли система этим характеристикам, можно проверить её по следующему списку вопросов.

- Пользователь понимает, что системой можно управлять по средствам жестов? Как пользователь это понимает?
- Визуальная модель учитывает скорость выполнения жеста?
- Понимает ли пользователь, что сейчас он выполняет компьютерный жест? Как пользователь это понимает?
- Системе удалось распознать жест? Как пользователь это понимает?
- Присутствует жест отмена? Как пользователь это понимает?
- Для пользователя доступен алфавит жестов?
- Возможно ли адаптировать жест "под себя"? И как быстро можно это сделать.

Это общие вопросы. Этот список необходимо расширить вопросами, в зависимости от типа системы, области применения и задачи.

Управление жестов должно быть в каком то смысле *инструментальным*. Когда мы используем молоток, мы не думаем, как нужно держать молоток, мы полностью сосредоточены на выполнение задачи - забить гвоздь. Молоток становится продолжением нас, мы расширяем свои физические способности с его помощью. Система жестов - это скрытое управление. Пользователь не должен даже догадываться о том, как он использует жесты для решения задачи. Они должны быть прозрачны, скрыты.

Чтобы проверить насколько прозрачно управление жестов, можно спросить пользователя ощущает ли он наличие специального управления или нет. Если нет, это очень здорово!

ПРОЕКТИРОВАНИЕ ЖЕСТОВ В РАМКАХ USER-CENTER DESIGN

Человеко-компьютерное взаимодействие это научная дисциплина, главным предметом которой является то, как люди используют компьютерные системы, как следует разрабатывать такие системы и как обеспечить более эффективное использование таких систем. Дисциплина объединяет такие области как психологи, социология, эргономика, информатика, дизайн и другие.

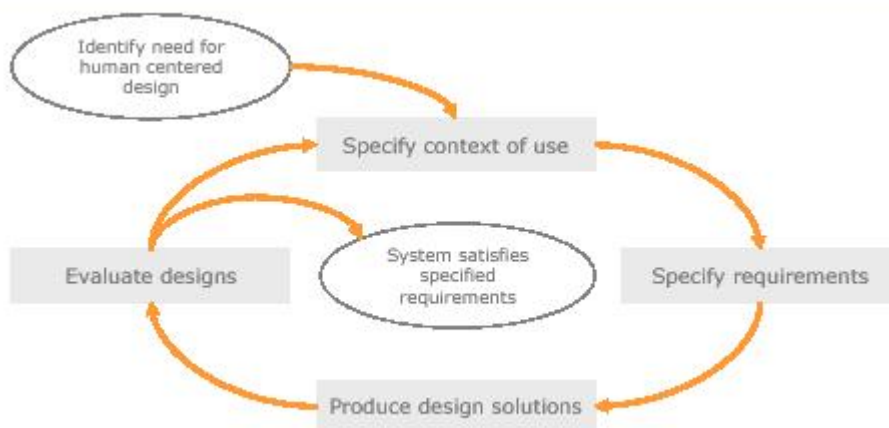
Основной подход, который используется в компьютерном взаимодействии основан на проектировании, основанном на нуждах пользователя (user-center design, UCD). Это подход к проектированию компьютерных систем, при котором главным является пользователи, их нужны и требования. При использовании данного подхода достигается высокий уровень юзабилити (англ. usability — дословно «возможность использования», «способность быть использованным», «полезность»). Международный стандарт ISO 9241-11 определяет юзабилити как *"степень, с которой продукт может быть использован определёнными пользователями при определённом контексте использования для достижения определённых целей с должной эффективностью, продуктивностью и удовлетворённостью"* (англ. "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use"). При этом относительная важность всех трёх аспектов определяется этим самым контекстом. Юзабилити это свойство системы или продукта. Если продукт обладает высокой степенью юзабилити (или продукт является юзабельным), это значит, что пользователи достигают своих целей при использовании этого продукта максимально эффективно и просто.

Существуют несколько общепризнанных стандартов по юзабилити, которые могут быть применимы в проектирование человеко-компьютерного взаимодействия. Некоторые из этих стандартов являются обязательными, некоторые носят рекомендательных характер. Существуют

несколько стандартов так или иначе относящийся к области юзабилити или проектированию человеко-компьютерного взаимодействия. Например, это:

- ISO 9241: Ergonomic requirements for office work with visual display terminals
- ISO 13407: Human-centred design processes for interactive systems
- ISO 18529: Human-centred lifecycle process descriptions
- ISO 14915: Software ergonomics for multimedia user interfaces
- ISO 16071: Software accessibility
- ISO 16982: Usability methods supporting human centred design

В предыдущей главе я описывала характеристики, которыми должна обладать система, основанная на жестовом взаимодействии. В этой мне бы хотелось уделить внимания вопросам, связанным с процессом проектирования таких систем. Существует международный стандарт, который является основой для многих методик UCD. Этот стандарт (ISO 13407: Human-centred design processes for interactive systems) определяет общий процесс для включения человеко-ориентированной деятельности всюду по жизненному циклу продукта, но не определяет точные методы.



В модели UCD определено 4 стадии, которые составляют основной цикл работы:

1. **Определить контекст использования.** Выявить пользователей, которые будут использовать продукт, зачем они будут его использовать и в каких условия они будут его использовать.

2. **Специфицировать требования.** Выявление бизнес-требований, пользовательских задач, которые должны быть осуществимы для успешности продукта.
3. **Разработать решение.** Эта часть процесса может быть сделано поэтапно, от концептуальной модели до конечной разработки.
4. **Произвести оценку.** Оценка - в идеале через юзабилити-тестирование с реальными пользователями - является наиболее важной частью процесса.

Процесс заканчивается и продукт может быть выпущен, после того как будут удовлетворены требования.

Методология UCD не специфицирует как либо процесс, однако UPA (Usability Professionals' Association) опубликовала постер, в котором показывает типичный процесс UCD. В этой версии деятельности UCD разбиты на четыре этапа: анализ, проектирование, внедрение и развертывание, а также предлагаются определенные деятельности для каждого этапа. К ним относятся:

- **Фаза анализа**

- Поговорить с основными заинтересованными сторонами; определить видение продукта;
- включить задачи по юзабилити в плане проекта;
- собрать междисциплинарную команду, чтобы провести разностороннюю экспертизу;
- разработать юзабилити цели и задачи;
- провести анализ аналогичных продуктов;
- разработать профили пользователей;
- провести анализ задач пользователей;
- разработать документ о сценариях пользователей;
- разработать документ о пользовательских требованиях.

- **Фаза проектирования**

- Провести коллективное обсуждение концепций дизайна и метафоры;
- разработать раскадровку экранов и навигационную модель;
- разработать пошаговые концепции дизайна;

- создайте бумажные прототипы с низкой точностью детализации;
 - проведите юзабилити-тестирование на полученном прототипе;
 - разработать более детальный дизайн;
 - провести юзабилити-тестирование еще раз;
 - разработать спецификацию дизайна.
- **Фаза разработки**
 - Проводить эвристическую оценку;
 - команда разработчиков и дизайнеров должна работать в тесном сотрудничестве;
 - проводить юзабилити-тестирование как можно раньше.
 - **Фаза развертывания**
 - Использовать различные исследования для получения обратной связи с пользователями;
 - провести специальные исследования для получения информации о фактическом использовании продукта;
 - проверить цели, используя юзабилити-тестирование.

Стоит обратить внимание, на то, что юзабилити-тестирование появляется несколько раз на протяжении всего процесса, начиная с первого этапа до последнего.

В своей работе я обсуждаю различные аспекты, связанные с проектированием систем на базе жестов. В этой главе мне бы хотелось поговорить о том, какие дополнительные деятельности нужно произвести на каждой фазе разработки, в рамках типичного процесса UCD.

Для **фазы анализа** в проектировании жестов управления я рекомендую сосредоточить внимание на главной задаче, в терминах деятельности, на контексте и среде использования и на технической реализуемости. Данные пункты отобраны, так как жест, это в первую очередь *действие*, это тот самый *глагол*, который должен *отсутствовать* в интерфейсах. Возникает некий парадокс: в итоге должен получиться такой интерфейс пользователя в котором нет глаголов действия, однако все глаголы присутствуют в действиях пользователя и осуществляются за счет поддержки контекста.

Техническая возможность надежно и точно распознать жест является в проектировании жестов приоритетной. Так как, если система будет часто ошибочно распознавать жест, или же наоборот, не распознавать в экстренной ситуации действия пользователя, то лучше, чтобы такой системы не было. Все риски, связанные с проблемами распознавания и осуществления возможности заменить какое-то действие на жест должны обсуждаться с командой разработчиков. Технические требования связанные с возможностью распознавания, должны быть вынесены в специальный документ (в дополнение к пользовательским требованиям).

Перед написанием технических требований к распознаванию жестов, необходимо проработать документ об среде и условиях использования системы под управлением жестов. Этот документ также обычно фигурирует при разработке традиционных интерфейсов, однако я его выделяю, так как он может повлиять на возможность распознавания жеста.

Чтобы получить список глаголов-действий, можно действовать точно также, как, когда мы получаем список объектов системы: из сценариев работы пользователя подчеркиваем все глаголы. Далее объединяем их в группы по близости смысла. В итоге у нас должен получиться список ключевых действий взаимодействия пользователя с системой. Этот список можно разделить на два списка: первый это универсальный, не зависящий от контекста использования, а второй - список действий с привязкой к контексту. Это удобно, так как поможет выделить некое основное меню управления системой. Особенность универсальных жестов системы в том, что они должны быть *уникальны* и *актуальны* (т.е. доступны для выполнения) в *любой* момент работы системы. Если какое-то действие не отвечает этому требованию, то скорее всего его нужно отнести ко второй группе. В первом списке у нас может (должен?) быть специальный жест *Отмена*. Этот жест должен отменять предыдущие действие. Уже наличие только этого жеста дает огромное преимущество интерфейсам под управлением жестов над традиционными интерфейсами.

Таким образом, в дополнение к "стандартным документам", необходимо иметь после стадии анализа:

- список универсальных, общих глаголов-действий;

- список глаголов-действий к привязке к контексту;
- условия использования или среда использования глаголов-действий;
- ограничения на распознавание жестов.

Как можно заменить тут не идет речи не о каком то возможном жесте.

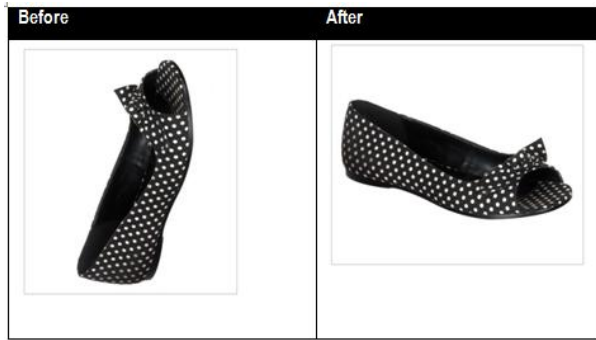
В фазе проектирования, перед разработкой прототипа с низкой детализацией, необходимо немного поработать со списком глаголов. В главе посвященной обзору жестов взаимодействия приводиться классификация жестов по функциональности: символические, дейктические, изобразительные, панто-мимические, акценты и связующие жесты. Необходимо связать каждый глагол из списка с классами функциональности жестов. Компьютерная метафора жеста должна быть основанная на группе "реальных" жестов. Далее нужно дополнить раскадровку экранов схематическими жестами. Под стрелочками-переходами между экранами изображаем глаголы-действия в виде жестов. Это изображение, может в дальнейшем стать прообразом обратной связи распознавания управляющего жеста. На самом деле, как уже говорилось, обратная связь не обязательно должна быть только графической. Если есть возможность, то можно использовать вибрацию, звук или что-то еще для информирования пользователя о том что жест распознан (и/или, если необходимо, что это за жест). Но это скорее исключение, чем правило, поэтому в большинстве случаев придется позаботиться о хорошей графической обратной связи.

Таким образом, в стадии проектирования мы должны:

- построить компьютерную метафору жеста, связывающую глагол-действие и функциональную группу "реальных" жестов;
- спроектировать обратную связь, позволяющую понять, что жест распознан и, если необходимо, что это за жест и какие жесты допустимы.

Иногда спроектировать жест не так просто, как бы нам хотелось. Можно использовать прием с карточками, на которых изображены две картинки: До (Before) и После (After). Требуется нарисовать и/или описать действия, которые нужно совершить, чтобы получить желаемое.

Рассмотрим этот способ на примере проектирования жеста-манипуляции. На рисунке представлена карточка для проектирования жеста. Довольно очевидно, какое действие мы должны совершить, чтобы



повернуть картинку. Этот способ не панацея, однако он хорошо подходит для проектирования жестов для сенсорных экранов. Если дополнить карточку визуальным представлением жестов, то можно использовать, скажем, при проектировании жестов на базе акселерометра.

Во время фазы проектирования юзабилити-тестирование используют для подтверждения выбора универсальных и контекстозависимых жестов, а также корректность отображение необходимой обратной связи о выполнении жеста и возможности выполнить действия в рамках сценария. Выходом из фазы проектирования будет спецификация дизайна, достаточно подробная для перехода в фазу разработки.

Когда мы приходим в **фазу разработки**, у нас на руках спецификация дизайна, в которой представлены зарисовки жестов управления. По сути эта стадия ни чем не отличается от привычной стадии разработки.

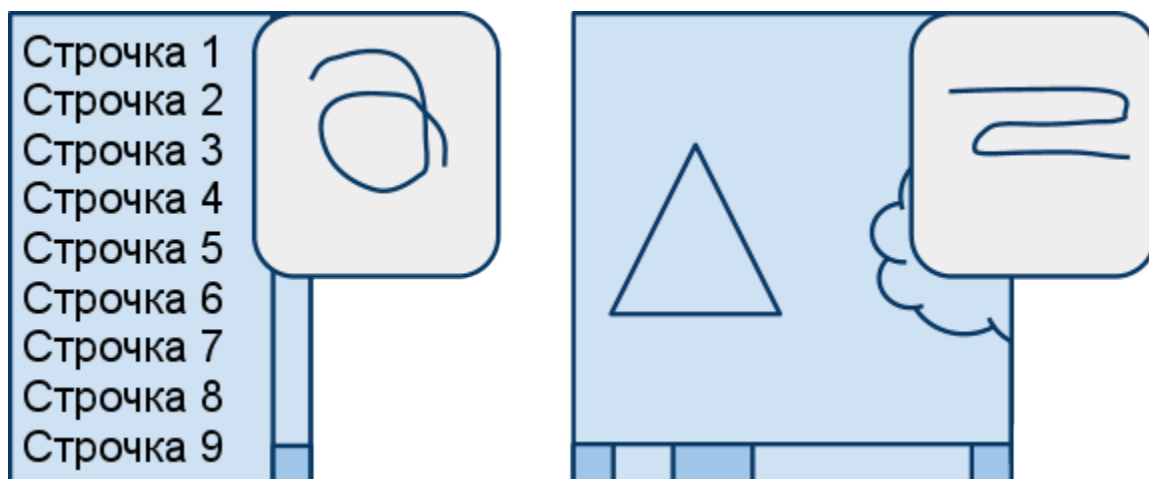
При проектировании жестов, в **фазе развертывания** мы может понять то, насколько просто и легко пользователи могут выполнить спроектированные жесты управления (если пользователь не заметил, что его действия это специальный компьютерный жест, то можно поздравить проектировщика=), насколько спроектированный жест доступен (если пользователь интуитивно догадался, что какой-то жест поддерживается системой, то можно поздравить проектировщика) и другие аспекты оценки, показывающие, что продукт обладает высокой степенью юзабилити.

ПРИЛОЖЕНИЕ

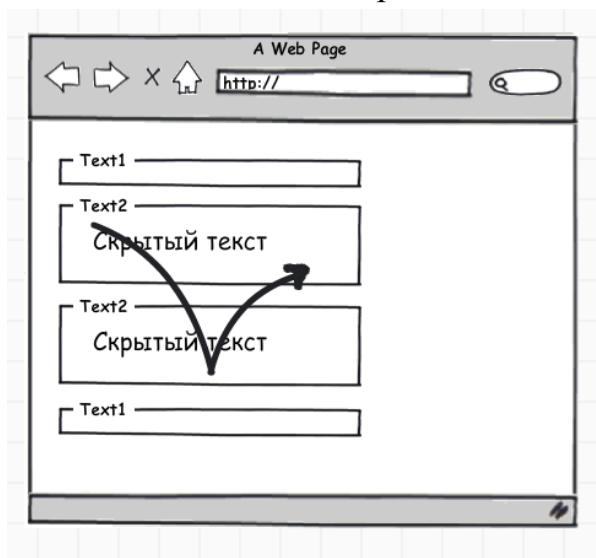
В ходе работы было разработано несколько приложений, которые так или иначе подтверждали (или опровергали) изложенные мысли.

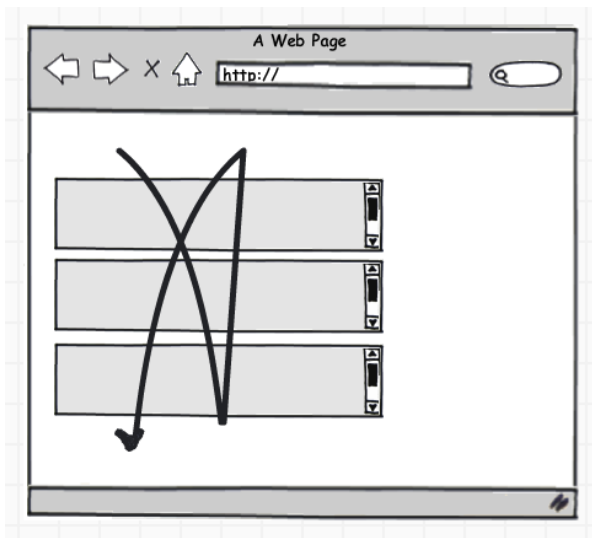
Примеры жестов для компьютерной мыши

Так, например, для проверки возможности создать на базе компьютерной мыши жесты управления, которые бы выгодно отличались от традиционного метода, но при этом использовали обычные элементы интерфейса, были разработаны специальные жесты для работы с большими списками. Например, такой элемент как линейка прокрутки, находится в противоречии с одним из основных принципов психологии восприятия: у человека может быть только одна точка активного внимания. При использовании же линейки прокрутки приходится смотреть в две совершенно различные точки – на прокручиваемое изображение (не пора ли остановиться) и на линейку. Если мы используем жесты, то посредник, в виде линейки исчезает. На рисунке ниже показан способ убрать линейку из интерфейса. В первом случае это вертикальный скроллинг, пользователь мышкой "рисует" круги на специальной зоне, а во втором случае, соответственно, это горизонтальный скроллинг, где пользователь "рисует" линии в сторону желаемого смещения. Этот способ одинаково хорошо работает и с компьютерной мышкой, и с сенсорной панелью (тачпад). Можно предположить что в телефонах с сенсорным экраном он также будет более удобен, если зеленую "активную" зону поместить прямо на список. Здесь и ниже, приводятся схематические рисунки. Сделано это для того, чтобы проще было передать суть и главные идеи.



Для больших списков жеста дают явное преимущество. Например, если мы используем графический элемент "гармошка", то открыть несколько соседних панелек, можно одним жестом, похожим на букву "V". И также быстро закрыть несколько, используя жест, похожий на букву "X". На рисунках ниже показаны схематические представления жестов.





Примеры для жестов с использованием мыши были реализованы на JavaScripts, что позволяет их использовать на WEB-сайтах. Эти и другие примеры были реализованы с использованием библиотеке jQuery (<http://jquery.com/>) и алгоритма распознавания, предложенным Andrew D. Wilson ($\$1$ GESTURE RECOGNIZER), реализованным на JavaScripts (<http://depts.washington.edu/aimgroup/proj/dollar/>).

Прототип системы управления презентацией

В ходе работы также был разработан прототип системы для управления презентацией с использованием акселерометра, в частности с помощью контроллера WiiRemote. Этот прототип интересен тем, что он выгодно использует поддержку контекста.

ЗАКЛЮЧЕНИЕ

ЛІТЕРАТУРА

1. Buxton, Bill. Sketching User Experience. Morgan Kaufmann Publisher, 2007.
2. Saffer, Dan. "Designing Gestural Interfaces". O'Reilly, 2009.
3. Gesture-Based Communication in Human-Computer Interaction. 5th International Workshop, GW 2003, Genova, Italy, April 15-17, 2003.
4. Norman, D. A. (2002). The Design of Everyday Things. New York: Basic Books.
5. Norman, D.A. & Draper, S.W. (Eds.) (1988). User Centered Systems Design, Hillsdale, N.J.: Lawrence Erlbaum Associates.
6. McNeill, D. (1985) So you think gestures are nonverbal? Psychological Review. 92, 350-371.
7. McNeill, D (1992) Hand and Mind: What Gestures Reveal about Thought. University of Chicago Press.
8. Kurtenbach, G. & Hulteen, E. (1990). Gestures in Human-Computer Communications. In B. Laurel (Ed.) The Art of Human Computer Interface Design. Addison-Wesley, 309-317.
9. Rime, B., Schiaratura, L. (1991). Gesture and speech. In R. S. Feldman, B. Rime, (Eds.), Fundamentals of Nonverbal Behavior (pp. 239-281). New York, NY: Cambridge University Press.
10. Lee, C. & Xu, Y. (1996). Online, Interactive Learning of Gestures for Human/Robot Interfaces. 1996 IEEE International Conference on Robotics and Automation, vol. 4, 2982-2987.
11. HUMAN-COMPUTER INTERACTION, 2000, Volume 15, pp. 263–322. "Designing the User Interface for Multimodal Speech and Pen-Based Gesture Applications: State-of-the-Art Systems and Future Research Directions".
12. Hauptmann, A.G. & McAvinney, P. (1993). Gestures with Speech for Graphics Manipulation. Intl. J. Man-Machine Studies, 38, 231-249.

13. Buxton, William. Human Input to Computer Systems: Theories, Techniques and Technology (<http://www.billbuxton.com/inputManuscript.html>)
14. Billingham, Mark. Gesture Driven Input (<http://www.billbuxton.com/input14.Gesture.pdf>)
15. Axel G.E. Mulder (1996). Hand gestures for HCI. Technical Report, NSERC Hand Centered Studies of Human Movement project. Burnaby, BC, Canada: Simon Fraser University.
16. Ekman P., Friesen W. V. The repertoire of nonverbal behavior: categories, origins, usage and coding // *Semiotica*. 1969, v. 1, p. 49–98.
17. Ekman P., Friesen W. V. Hand movement // *J. Communication*, 1972, v. 22, p. 353–374.
18. Shneiderman, B. (1983). Direct Manipulation: a step beyond programming languages, *IEEE Computer*, 16(8), 57-69.
19. Shneiderman, B. (1991). Touch screens now offer compelling uses. *IEEE Software*, 8(2), 93-107.
20. Pierre David Wellner. Interacting with paper on the DigitalDesk, 1994.
21. M. Krueger. Videoplace (1974)// <http://design.cca.edu/graduate/uploads/pdf/videoplace.pdf>.
22. The Sensor Frame Graphic Manipulator NASA Phase II Final Report// http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19940003261_1994003261.pdf
23. <http://www.elotouch.com/AboutElo/History>
24. <http://www.xrce.xerox.com/Research-Development/Historical-projects/DigitalDesk-to-CamWorks>
25. <http://oblong.com>
26. <http://zig.media.mit.edu>