

Курс «Введение в естественно-интуитивное взаимодействие с компьютером»

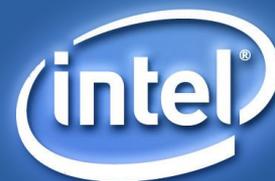
Лекция 1

Новые возможности в человеко-машинном интерфейсе



Содержание лекции

- Введение
- История развития человеко-компьютерного взаимодействия
- Что такое естественно-интуитивное взаимодействие
- Техническое обеспечение
- Сравнение современных камер
- Средства разработки: Intel Perceptual Computing SDK
- Области применений и примеры приложений
- Выводы
- Контрольные вопросы



Введение (1)

- Электронно-вычислительные машины заняли место рядом с человеком с середины XX века, и со временем данное взаимодействие только увеличивается.
- На заре развития электронно-вычислительной техники общаться с компьютерами могли только «посвященные».



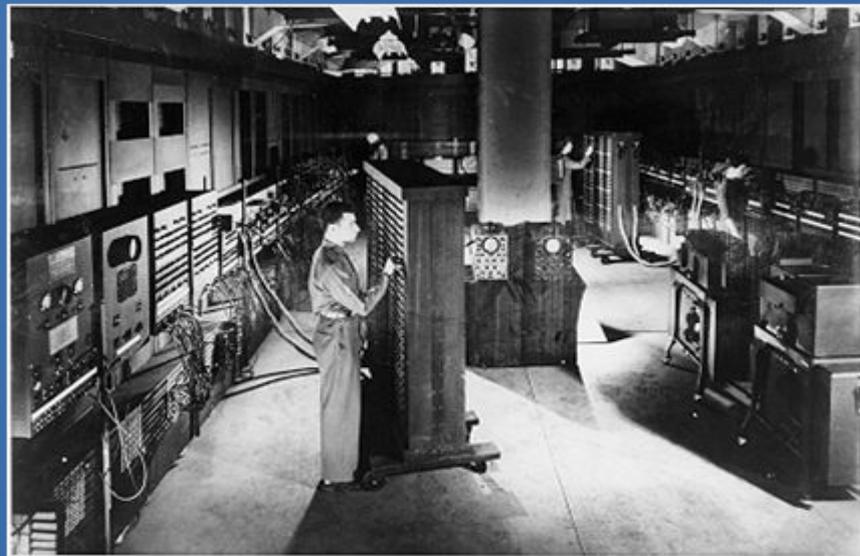
Введение (2)

- Разработчики стремились сделать взаимодействие с вычислительными устройствами как можно более естественным для человека. Однако развитие этого направления сдерживалось недостаточным технологическим уровнем.
- В последнее время в деле *гуманизации пользовательских интерфейсов* наметились серьезные продвижения. К примеру, появление сенсорных экранов и функций множественного касания позволяет взаимодействовать компьютером на интуитивном уровне.



История развития человеко-компьютерного взаимодействия (1)

- Середина XX века - первые электронно-вычислительные машины были ориентированы на выполнение трудоемких расчетов и на взаимодействие со специалистами.

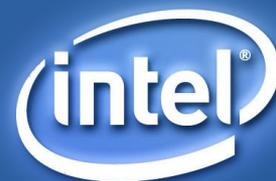


Взаимодействие человека и ЭВМ
ENIAC

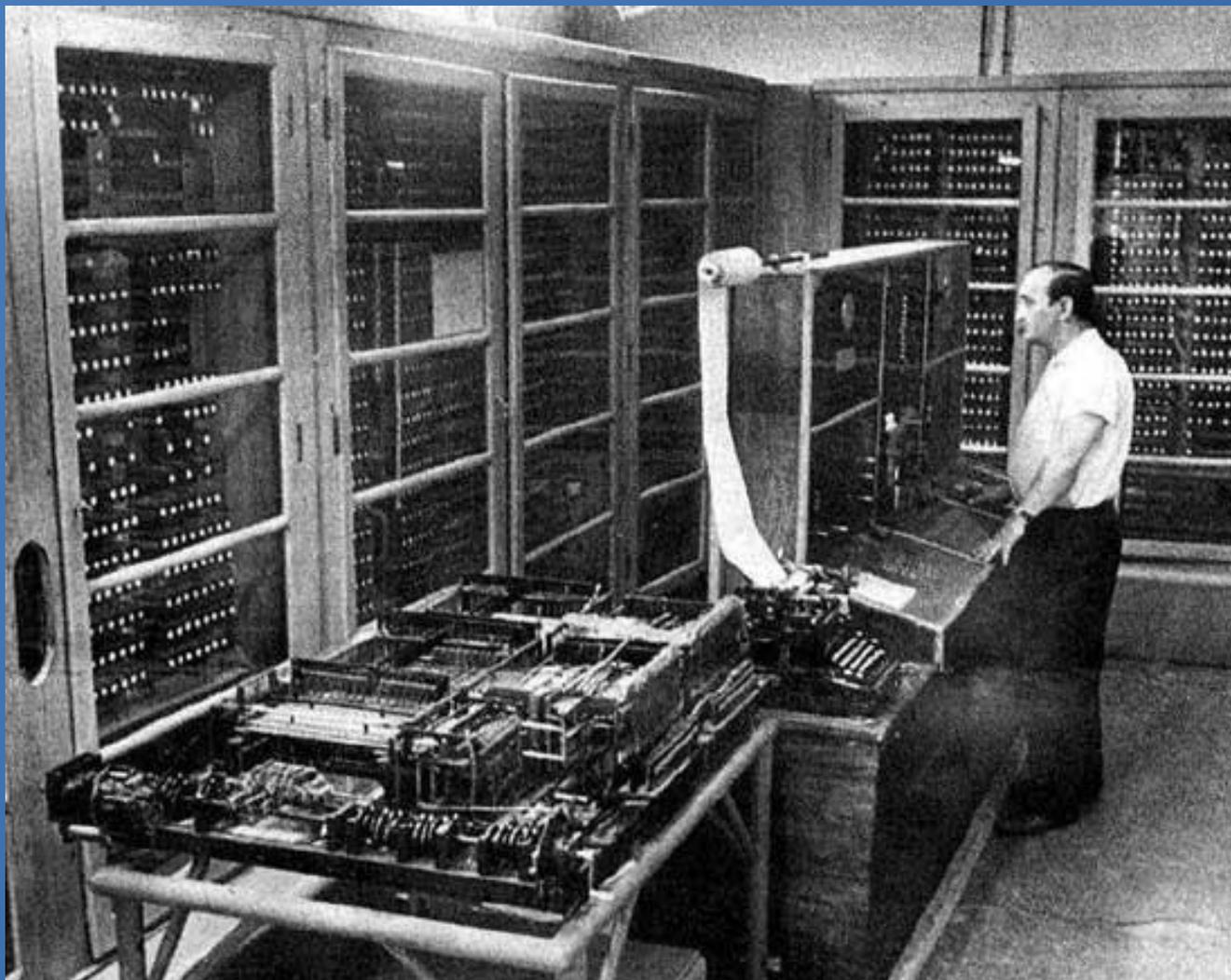


История развития человеко-компьютерного взаимодействия (2)

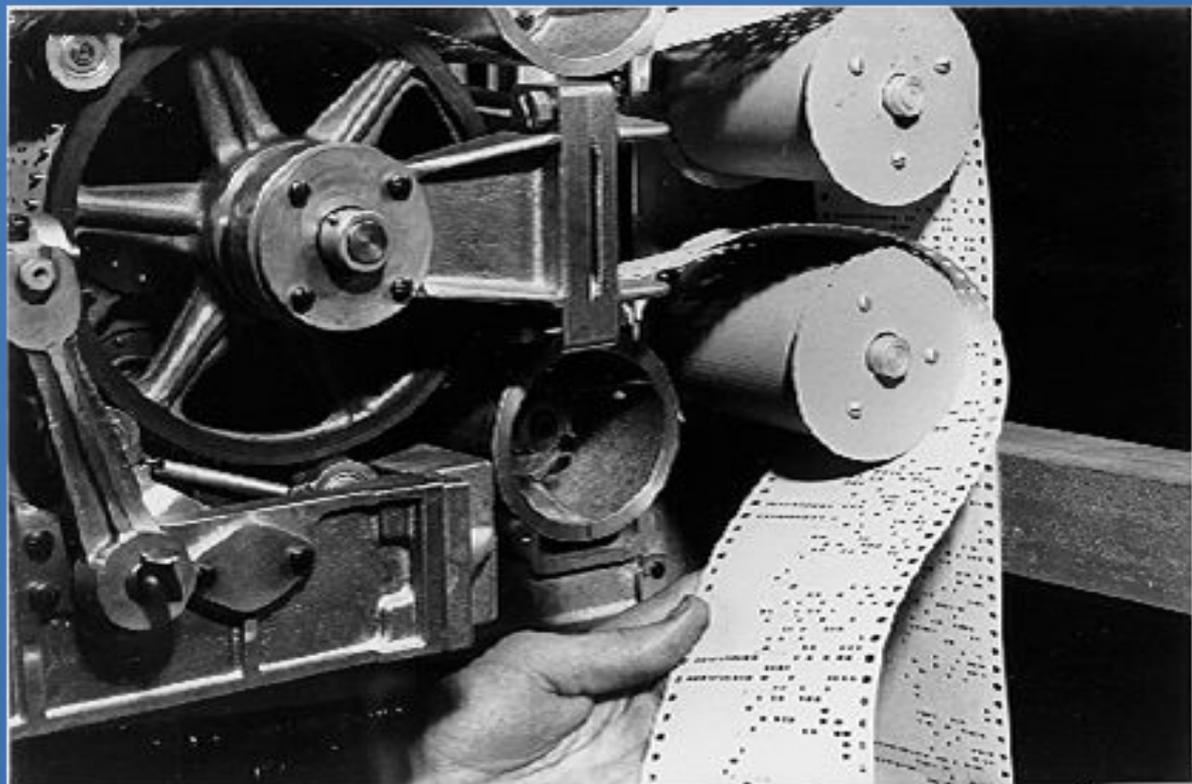
ENIAC (Electronic Numerical Integrator And Calculator) проект 1943-46 гг. под руководством Дж. Моучли и Дж. Эккера (Пенсильванский университет). Группа разработчиков – 200 чел. Цель – автоматизация расчетов для составления баллистических таблиц различных видов оружия (заказ баллистической исследовательской лаборатории Армии США). Вес 30 тон 18 тыс. радиоламп, 10 тыс. конденсаторов, 6 тыс. переключателей, 500 тыс. паяных соединений Мощность 150 киловатт (~1000 TV) Площадь 150 м² Скорость 5000 операций В среднем 1 лампа заменялась через 20 часов



Из истории развития ЭВМ

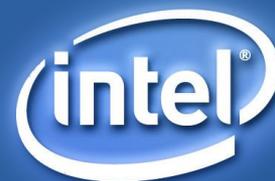


Из истории развития ЭВМ



Устройство чтения перфолент ЭВМ Mark I

(From Harvard University Cruft Photo Laboratory)



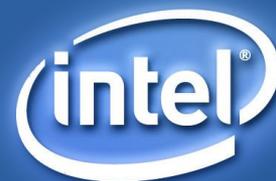
Из истории развития ЭВМ

```
A:dir
COMMAND COM 4896 8-23-83 1:15a
FORMAT COM 2688 1-01-80 1:01a
RECV EXE 1024 8-23-83 1:02a
DEBUG COM 6016 8-22-83 3:05p
CHKDSK COM 1728 8-22-83 3:00p
FILCOM COM 8320 8-22-83 3:03p
EDLIN COM 2432 8-22-83 3:06p
LINK EXE 41856 8-22-83 3:13p
EXEZBIN EXE 1280 8-22-83 3:07p
MASH EXE 70784 8-22-83 3:21p
SYS COM 608 8-22-83 3:23p
FORMAT OBJ 4224 8-22-83 3:25p
CREF EXE 13824 8-22-83 3:02p
LIB EXE 32128 9-20-83 2:18p
RDCPM BAK 1920 9-20-83 2:19p
RDCPM COM 9600 9-20-83 2:20p
RDCPM OBJ 132 1-01-80 1:04a
17 File(s)
A:
COPY 1 SKP 1 COPY TO SKP TO COPY LN KILL CANCEL BCKSPC
```



История развития человеко-компьютерного взаимодействия (2)

- 1960 год - Дж.К.Р. Ликлайдер (J.R.Licklider) выдвинул идею «симбиоза человека и компьютера» – объединения человеческого интеллекта и вычислительной техники для управления информацией. Были предложены промежуточные цели, достижение которых предполагает реализацию данной идеи.



История развития человеко-компьютерного взаимодействия (3)

Ближайшие цели

- разделение времени компьютера между пользователями;
- электронный ввод/вывод символьной и графической информации;
- интерактивные системы реального времени для обработки информации и программирования;
- крупномасштабные системы хранения и поиска информации.

Среднесрочные цели

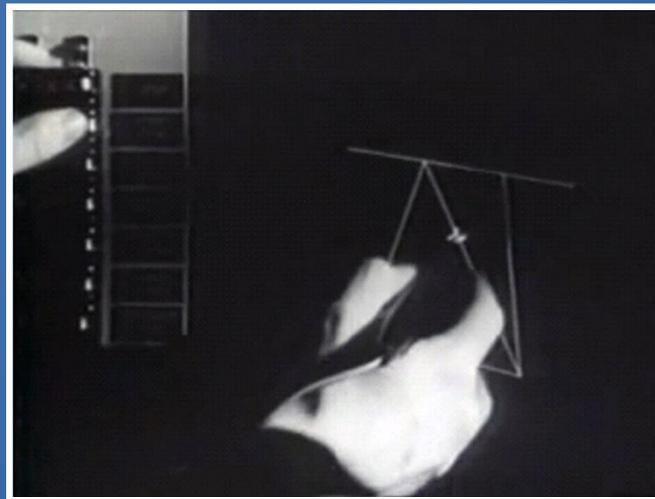
- координация объединения разработчиков для проектирования и программирования больших систем;
- способность ЭВМ распознавать речь оператора;
- способность ЭВМ распознавать рукописные тексты;
- возможность использования светового пера, в качестве устройства ввода координат и указки (световое перо – светочувствительное устройство, позволяющее выбрать точку экрана дисплея, указывая на нее).

Долгосрочные цели

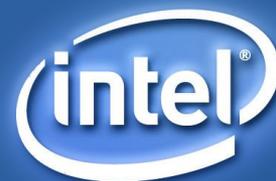
- понимание ЭВМ естественного языка;
- способность ЭВМ распознавать речь произвольного пользователя;
- эвристическое программирование.

История развития человеко-компьютерного взаимодействия (4)

- 1963 год - Айвен Сазерленд (Ivan Sutherland) разработал SketchPad – графический комплекс, прообраз будущих САПР, оказавший огромное влияние на формирование базовых принципов графических пользовательских интерфейсов.



SketchPad



История развития человеко-компьютерного взаимодействия (5)

Середина 60-х годов - командой Дугласа Энгельбарта была разработана среда NLS (oN-LineSystem), включающая в себя:

- принципиально новую операционную систему;
- универсальный язык программирования;
- электронную почту;
- разделенные экраны телеконференции;
- систему контекстной помощи;
- представлен прототип **WIMP-интерфейса** (windows, icons, menus, pointers).



История развития человеко-компьютерного взаимодействия (6)

Как побочный эффект проекта NLS был изобретен первый манипулятор типа мышь.



Первая компьютерная мышь (1964 год)

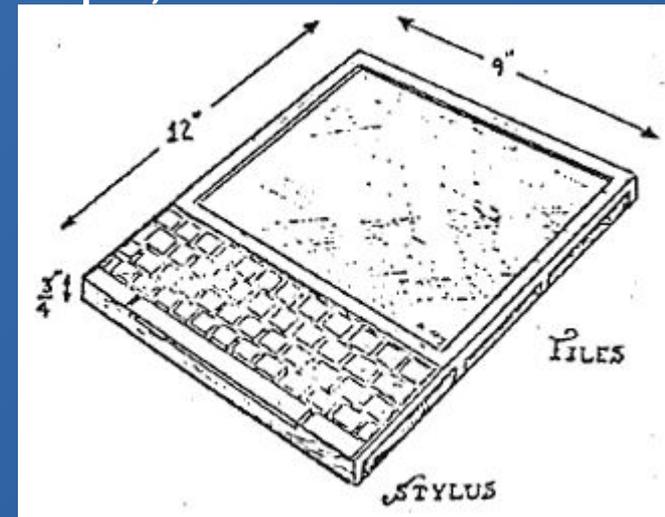
История развития человеко-компьютерного взаимодействия (7)



В 1969 году американский математик Алан Кей разработал принципы создания персонального компьютера.

С 1971 года Алан Кей занимался теоретической разработкой прототипа персонального компьютера, названного им Dynabook.

Концепция Dynabook описывала то, что сейчас известно как ноутбук, или планшетный ПК.



История развития человеко-компьютерного взаимодействия (8)

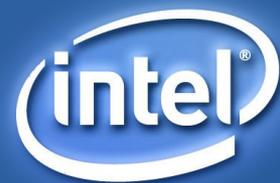
Несмотря на то, что графический интерфейс был описан еще в начале 70-х годов, в реальности взаимодействие пользователя с ЭВМ обеспечивалось за счет интерфейса командной строки (CLI, Command Line Interface).

К концу 70-х годов накопились технологии, позволяющие реализовать эргономическое проектирование вычислительной техники. При создании персональных компьютеров учитывалось удобство пользователей.



Человеко-компьютерное взаимодействие

- Человеко-компьютерное взаимодействие (HCI, Human-Computer Interaction) – это дисциплина, имеющая дело с проектированием, оцениванием и реализацией интерактивных вычислительных систем для использования человеком, а также с изучением основных явлений, связанных с этими вопросами.



Эволюция устройств ввода/вывода

	Устройства ввода	Устройства вывода
На заре эры вычислительной техники	<ul style="list-style-type: none">- переключатели и гибкие кабели со штекерами;- перфолента и перфокарты;- клавиатура.	<ul style="list-style-type: none">- экран дисплея;- телетайп;- принтеры.
Недавнее прошлое	<ul style="list-style-type: none">- клавиатура;- световое перо;- джойстик;- мышь;- микрофон.	<ul style="list-style-type: none">- прокручиваемый «стеклянный телетайп»;- текстовый терминал;- аудио.
Наши дни	<ul style="list-style-type: none">- цифровые перчатки и костюмы;- естественный язык;- интерактивная камера.	<ul style="list-style-type: none">- наголовный дисплей;- повсеместные вычисления;- автономные агенты.
Ближайшее будущее	???	???



Естественно-интуитивное взаимодействие

- Естественно-интуитивное взаимодействие - это эволюция в области взаимодействия с компьютером через естественные для человека способы и интерфейсы.
- Это не просто замена ввода с клавиатуры или тачскрина, это новый опыт во взаимодействии с вычислительной техникой.
- На сегодняшний момент у нас есть достаточно вычислительных мощностей для того, чтобы объяснить компьютеру то, что хочет от него пользователь, заставить его предугадывать намерения пользователя.



Техническое обеспечение (1)

Многие важные инженерные задачи должны быть решены с помощью:

- интеграции новых усовершенствованных датчиков в персональные компьютеры, ультрабуки, планшеты и смартфоны;
- стандартизованности и миниатюризации датчиков;
- камер, способных оценить глубину;
- микрофонов, определяющих направленность звука;
- тачпадов, чувствительных к давлению.



Техническое обеспечение (2)



CREATIVE

Interactive Gesture Camera

Особенности:

- небольшие размеры;
- работа на небольшом расстоянии;
- подходит для распознавания жестов, идентификации и трэкинга объектов, рук, лица, их синтеза, распознавания голоса.

Характеристики камеры:

- * Разрешение RGB сенсора: 720p (1280 x 720)
- * Разрешение IR Depth сенсора: QVGA (320 x 240)
- * Угол обзора (градусы): 73
- * Частота работы (кадров в секунду): 30
- * Расстояние: от 15 см до 1 метра
- * Питание: USB 2.0 (< 2.5 Ватт)



Техническое обеспечение (3)



Камера
Kinect

Характеристики камеры Kinect:

* 3 камеры:

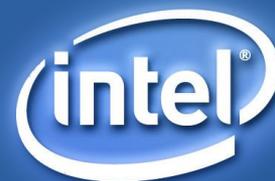
1 - цветная MT9M112, 1280x1024, 15 кадров в сек (30 кадров при 640x512);

2 - цветная MT9v112, 640x480, 30 кадров в секунду;

3 - ИК сенсор глубины MT9M001, инфракрасный черно-белый 1280x1024, 30 кадров в секунду;

* Расстояние от 1,2 метра до 3 метров

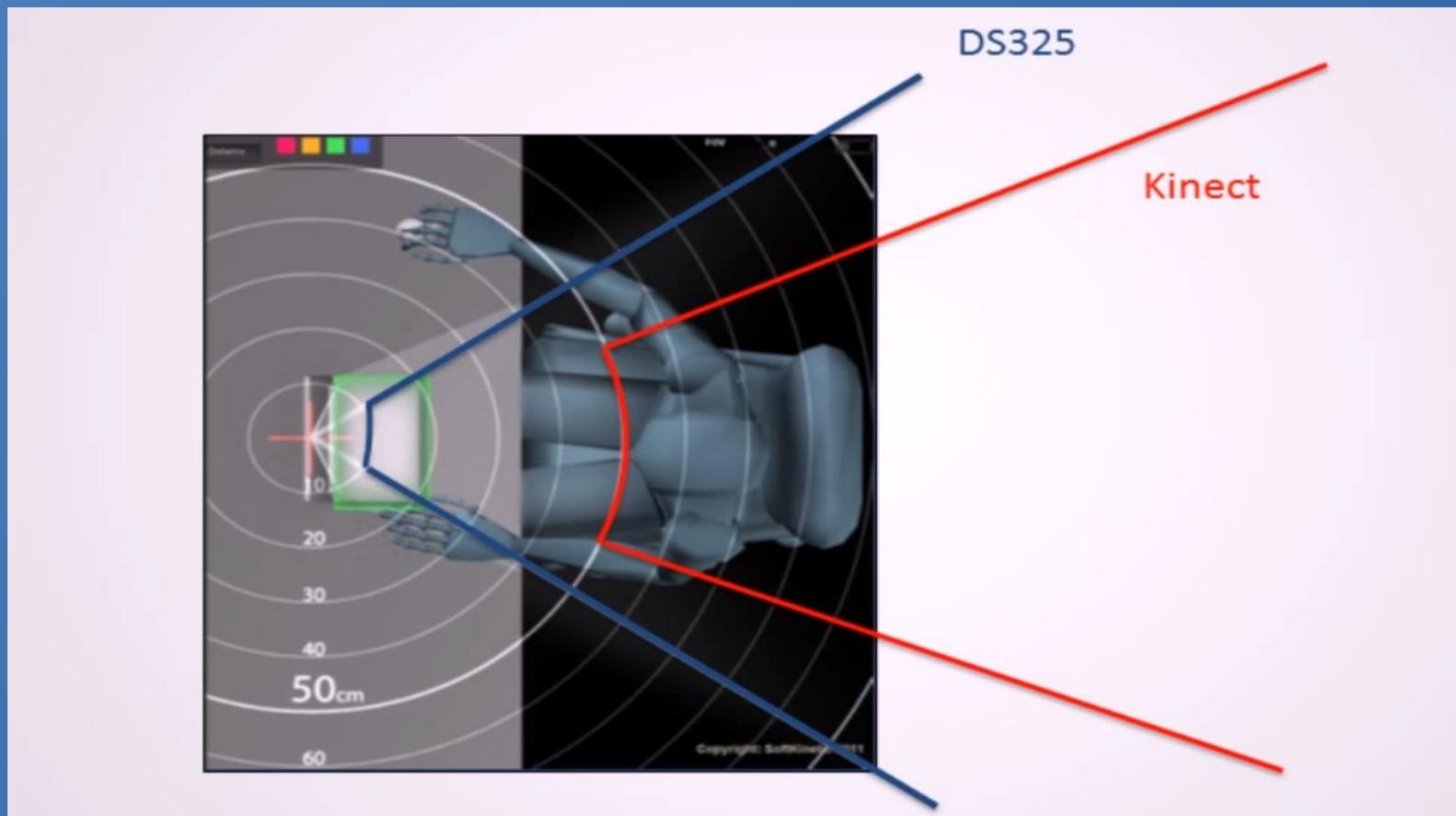
- Работа на больших дистанциях;
- Идентификация и распознавание жестов и тела человека, определение его в окружающей среде дома;
- Одновременное наблюдение за 6 людьми.



Техническое обеспечение (4)

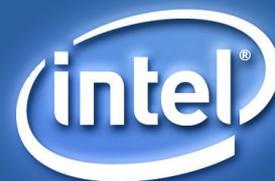


Техническое обеспечение (5)



Техническое обеспечение (6)

Характеристики	CREATIVE	Kinect
Минимальная дистанция (см)	15	120
Максимальная дистанция (см)	100	300
Горизонтальный угол обзора	72°	57°
Вертикальный угол обзора	54°	43°



Средства программной разработки

- Intel Perceptual Computing SDK
(<http://www.intel.com/software/perceptual/>)
- OpenCV (<http://opencv.org/>)
- CCV (<http://libccv.org/>)
- PCL (<http://www.pointclouds.org/>)



Intel Perceptual Computing SDK

- SDK предназначен для того, чтобы снизить стоимость разработки программных продуктов, в которых в качестве средств взаимодействия с пользователем необходимо использовать что-то отличное от клавиатуры или мышки



Особенности SDK

- Приложения могут получить доступ к функциональности SDK напрямую за счет стандартизации интерфейсов модулей ввода/вывода и алгоритмов;
- SDK предоставляет механизм для поиска конкретной реализации из множества доступных модулей;
- Реализует функции синхронизации выполнения и взаимодействие с другими библиотеками и фреймворками;
- SDK предоставляет набор вспомогательных классов для общих случаев использования (исходный код или библиотеки форм).



Модули обработки Intel Perceptual Computing SDK

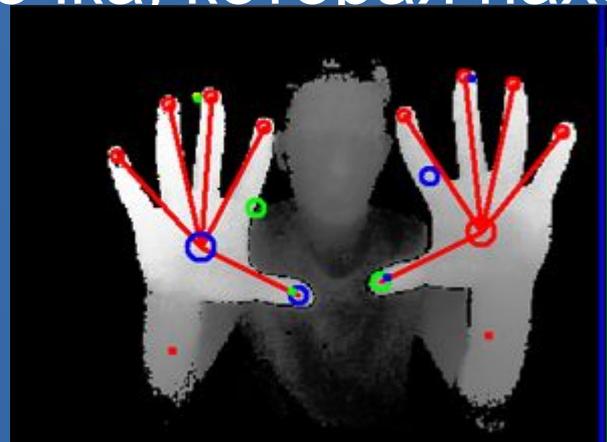
- Модуль распознавания жестов;
- Модуль распознавания лиц;
- Модуль голосового управления;
- Модуль отслеживания двумерных и трехмерных объектов;



Модуль распознавания жестов (1)

Отслеживания позиций пальцев на руке (7 точек):

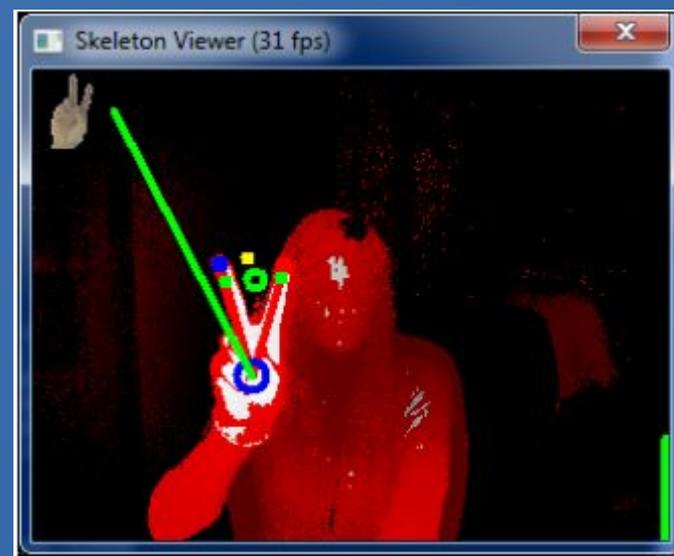
- Кончики пальцев;
- Центр ладони;
- Максимально видимая точка, которая находится у локтя.



Модуль распознавания жестов (2)

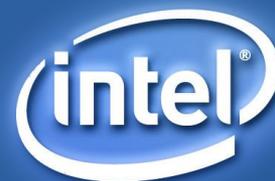
Распознавание стандартных жестов:

- Большой палец вверх;
- Большой палец вниз;
- Символ V (победа).
- Раскрытая ладонь.



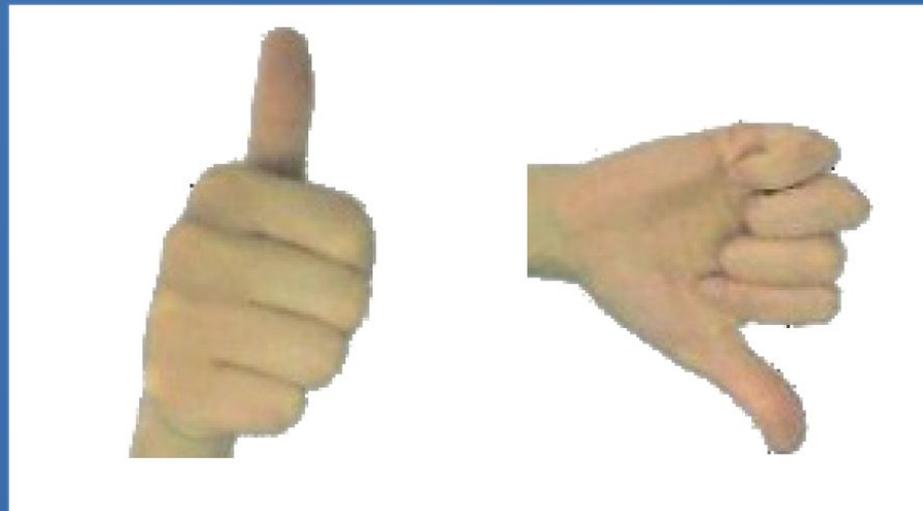
Распознавание динамических жестов:

- Взмахи руки влево, вправо, вверх, вниз, круговое движение, помахивание.



Основные распознаваемые позы (1)

- Открытость ладони
- «Большой палец вверх» или «Большой палец вниз»



Основные распознаваемые позы (2)

- Знак победы

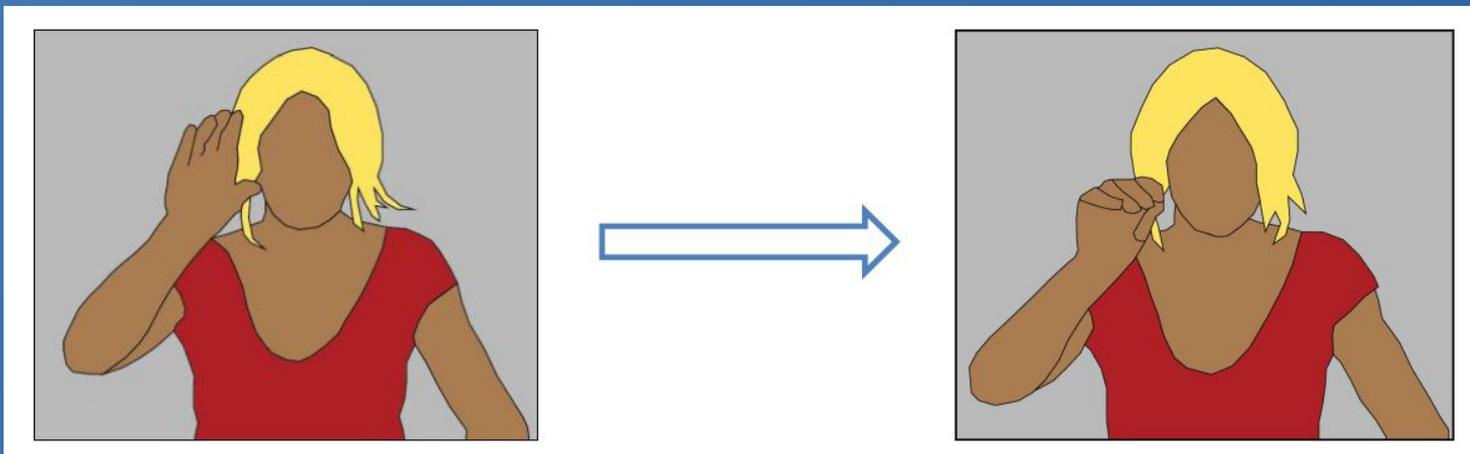


- Большая пятерка



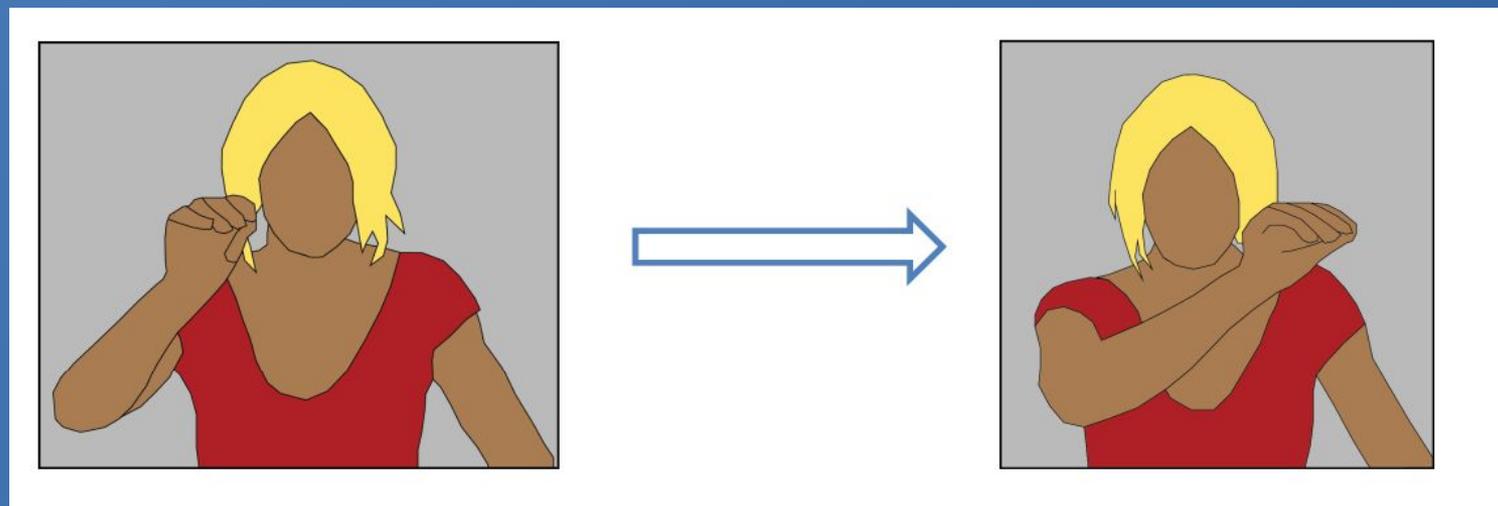
Основные распознаваемые жесты (1)

- «Захват» и «Сброс» - Перед началом выполнения этого жеста пользователь должен принять позу, в которой его большой палец и другие пальцы разъединены, следующим жестом все пальцы соединяются вместе в позу захвата. Обратные действия - разъединение пальцев отпускают объект.



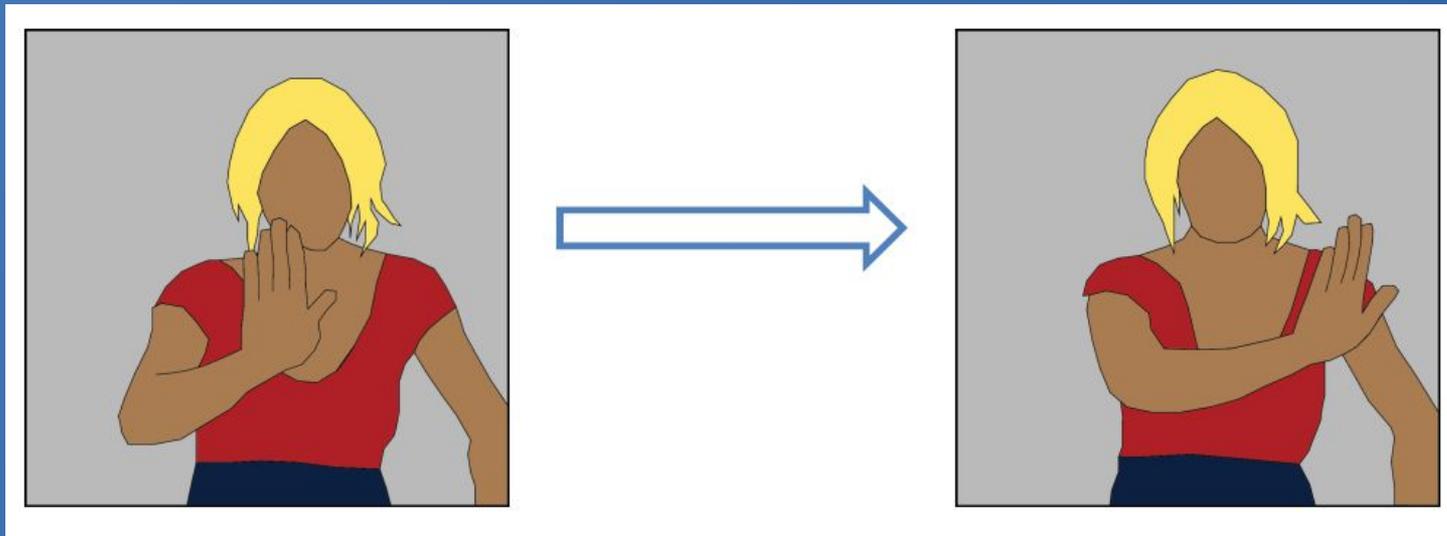
Основные распознаваемые жесты (2)

- «Движение» - После захвата объекта пользователь передвигает руку для перемещения объекта.



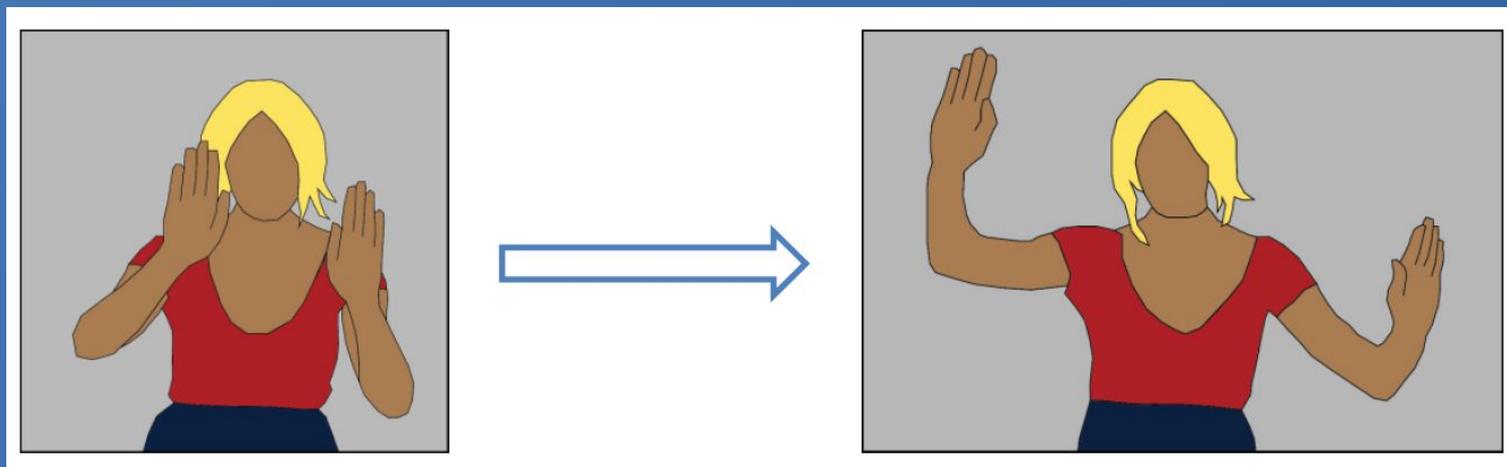
Основные распознаваемые жесты (3)

- «Панорамирование» выполняется только полностью прямой ладонью. Движение прямой ладонью панорамирует объект, но как только ладонь принимает расслабленную, слегка скрученную позу панорамирование останавливается.



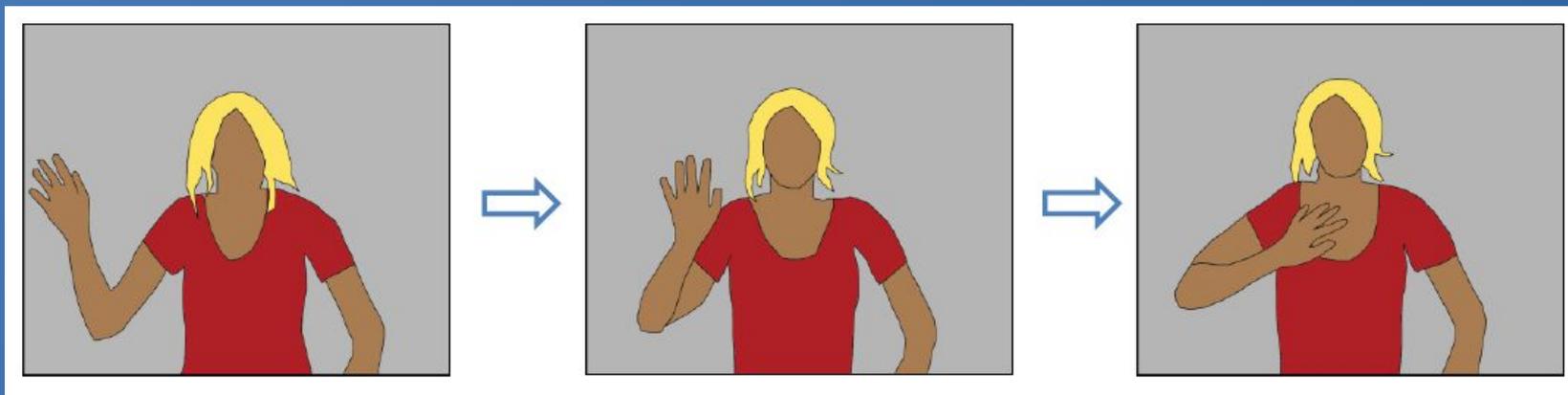
Основные распознаваемые жесты (4)

- «Масштабирование» выполняется за счет изменения расстояния между двумя ладонями. Масштабирование требует действия, которое завершит масштабирование, в противном случае пользователь не сможет его завершить без изменения масштаба.



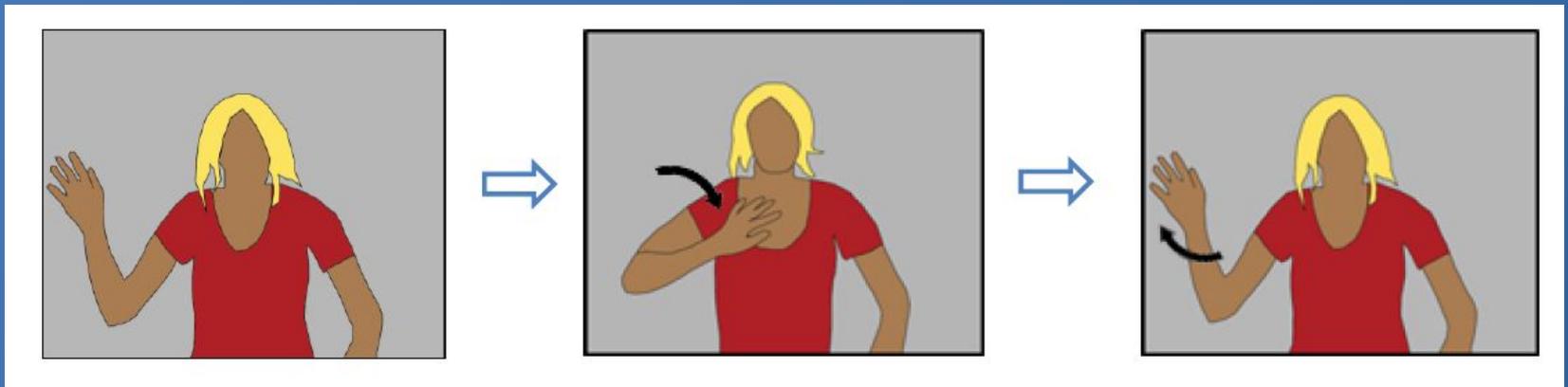
Основные распознаваемые жесты (5)

- «Помахивание» - При этом жесте, пользователь быстро машет рукой. Жест применяется для сброса, выхода из режима, либо для передвижения вверх по иерархии меню.



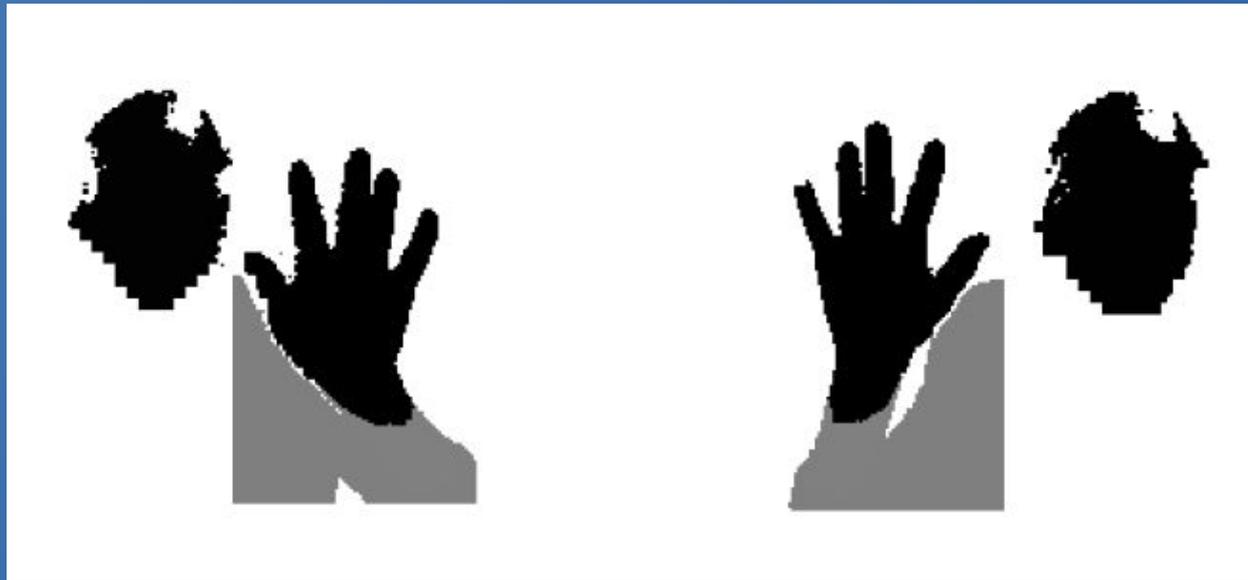
Основные распознаваемые жесты (6)

- «Круг» - Жест круг выполняется, когда пользователь соединяет все пальцы и двигает рукой по кругу



Основные распознаваемые жесты (7)

- Все жесты могут быть выполнены как правой, так или левой рукой, в жестах, выполняемых двумя руками (например, захват объекта двумя руками для изменения размера), не имеет значения, какая рука будет захватывать объект первой.



Основные распознаваемые жесты (8)

- Для многих жестов, число пальцев участвующих в жесте не имеет значения.



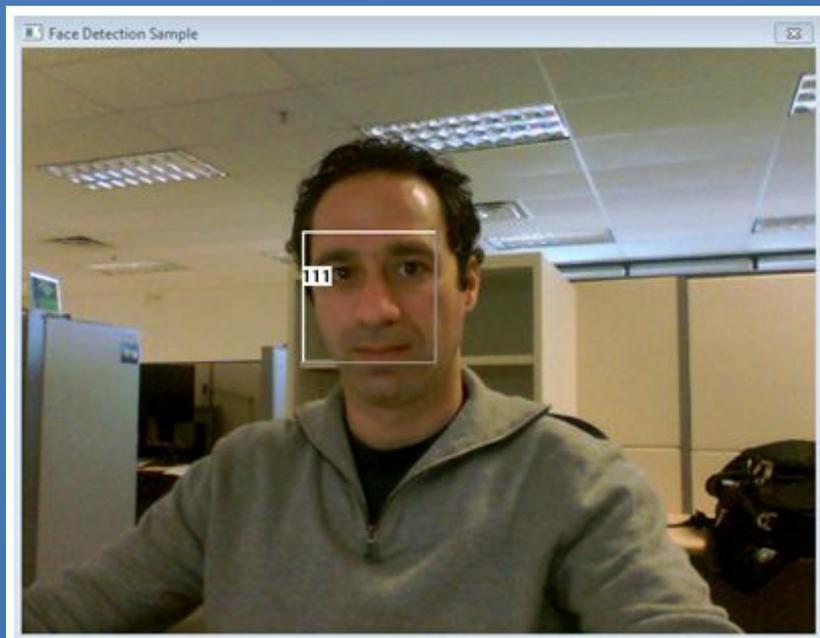
Модуль распознавания лиц (1)

Модуль позволяет получить большое количество информации о распознанном лице (7 точек на лице):

- Уголки глаз;
- Уголки рта;
- Кончик носа.



Модуль распознавания лиц (2)



Обнаружение
лица



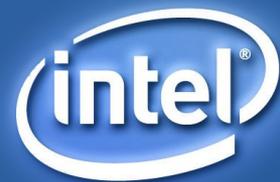
Распознавание
лица

Модуль распознавания лиц (3)

Модуль позволяет получить некоторую аналитическую информацию о распознанном лице:

- Возрастная группа лица присутствующего на картинке (ребенок, пожилой человек);
- Пол;
- Подмигивание;
- Улыбка.

Еще одна полезная особенность этого модуля заключается в том, что он может находить похожие лица.



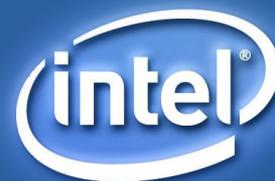
Модуль голосового управления

Модуль предлагает следующую функциональность:

- Реализация голосовых меню;
- Надиктовывание;
- Синтез речи.

На текущий момент доступен только английский словарь. Другие языки будут добавляться по мере сотрудничества с разработчиками голосового движка.

Сейчас для реализации голосового управления используется голосовой движок Nuance Dragon Assistant. Русский язык также скоро будет доступен.



Модуль отслеживания двумерных и трехмерных объектов (1)

Функционал данного модуля позволяет отслеживать плоские двумерные объекты. Достаточно создать модель объекта на компьютере и отслеживать его.

В результате будут получены:

- Параметры позиции объекта;
- Угол наклона и т.д.

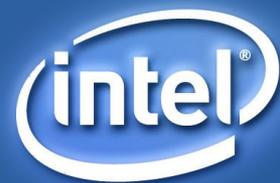
Также поддерживается отслеживание трехмерных объектов, которые задаются в виде модели (.obj файла).



Интеграция с игровыми движками

SDK реализует несколько уровней интерфейсов, которые позволяют сразу же использовать игровые движки. На текущий момент поддерживаются следующие игровые движки:

- Unity;
- Processing (использующий Java);
- openFrameworks.

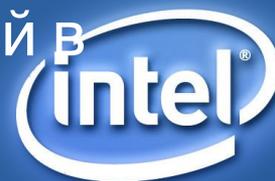


Области применения и примеры приложений (1)

Использование функций естественно-интуитивного взаимодействия



- взаимодействие игрока с трехмерным миром и другими игроками;
- использование голоса или жестов для манипуляций в игре.

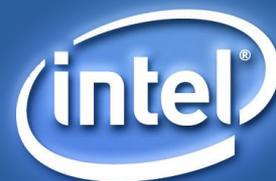


Области применения и примеры приложений (2)

- Распознавание пальцев на руках;
- Моделирование системы захвата какого-либо трехмерного
- Перемещение объекта;
- Указывание на данный объект.



«Железный
человек»
режиссёр Джон
Фавро



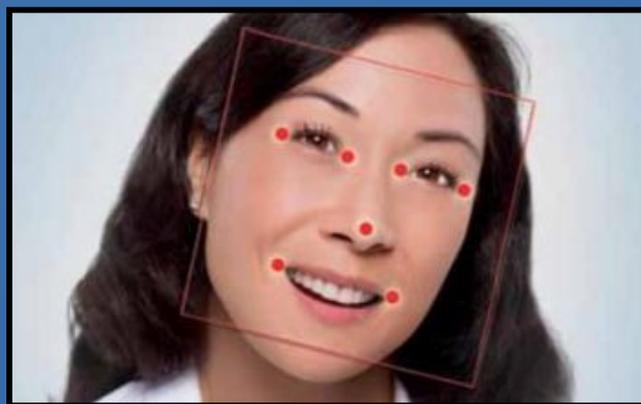
Области применения и примеры приложений (3)

- Распознавание движения рук или жестов может быть интерпретировано компьютером и использовано для взаимодействия. Могут быть реализованы такие жесты как перелистывание в разных направлениях (к примеру, переключать треки в плейлисте или перелистывать страницы сайта)



Области применения и примеры приложений (4)

- Функции распознавания лица и контрольных точек на лице позволят определить, где расположены глаза, нос или рот. Присутствует функция распознавания моргания, улыбки, распознавания пола и возрастной группы.



Области применения и примеры приложений (5)

- Простая идентификация личности (фейс-логин).
- Комбинация параметров лица, голоса и жестов для более точной идентификации.



Области применения и примеры приложений (6)

- Сегментация, основанная на глубине изображения, позволит разделить передний и задний фон.



К примеру, чтобы добавить красок виртуальному общению, задний фон может быть заменен каким-либо изображением, к примеру лунным ландшафтом или пляжем.

Области применения и примеры приложений (7)

Распознавание трехмерных объектов реального мира и комбинирование их с интерактивным контентом.

Трехмерная камера способна:

- распознать модель реального объекта
- добавить виртуальные графические и звуковые спецэффекты к объекту.

Таким образом, распознав группу объектов, положение отдельных объектов относительно других и их поведение, создается новая модель взаимодействия объектов.



Выводы (1)

Реализация функций естественно-интуитивного взаимодействия в интерактивных приложениях позволит:

- Отражать реальность, а не создавать копию реальности;
- Буквально, а не абстрактно отображать объекты реального мира;
- Создавать приложения интуитивно понятными;
- Создавать надежные приложения;
- Расширять функционал интерактивных приложений.



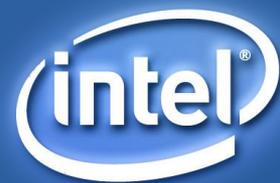
Выводы (2)

- Новый подход к управлению в играх;
- Идентификация личности в системах безопасности;
- Сегментация объектов переднего и заднего плана;
- Распознавание трехмерных объектов реального мира и комбинирование их с интерактивным контентом;
- И т.д.



Выводы (3)

С развитием технологий обработки речи, изображений и видео, взаимодействие человека с компьютером выходит на новый этап. В последние годы человеко-компьютерное взаимодействие было расширено и конечной целью является то, что связь между людьми и машинами должна стать похожа на коммуникацию человека с человеком.



Контрольные вопросы

1. Что включает в себя понятие человеко-машинный интерфейс?
2. Перечислите основные средства человеко-машинного интерфейса.
3. Где используются сенсорные дисплеи?
4. Какие виды клавиатур Вы знаете?
5. Перечислите сферы применения трекболов, джойстиков и манипуляторов типа «мышь».
6. В чем состоит принцип работы шлема виртуальной реальности?
7. Имеются ли принципиально новые, прорывные разработки в области создания средств человеко-машинного интерфейса? Охарактеризуйте эти разработки.
8. Существуют ли на сегодняшний день эффективные системы распознавания речи? В чем недостатки систем распознавания речи?

