

**Применение возможностей Intel
Perceptual Computing SDK для
расширения границ взаимодействия
людей с ограниченными
возможностями с внешним миром**

Лекция 2



Содержание лекции

- Введение
- Основы распознавания образов
- Основы распознавания жестов
- Intel Perceptual Computing SDK расширения границ взаимодействия людей с ограниченными возможностями с внешним миром
- Заключение
- Контрольные вопросы



Введение (1)

- В современном мире огромная роль отводится информации
- Людям приходится участвовать в непрерывном информационном обмене
- Необходимость быстрого получения доступа к информации, её обработке, представлении результатов анализа информации.



Введение (2)

«Информационная доступность» означает возможность беспрепятственного получения информации вне зависимости от особенностей потребностей пользователя.



Введение (3)

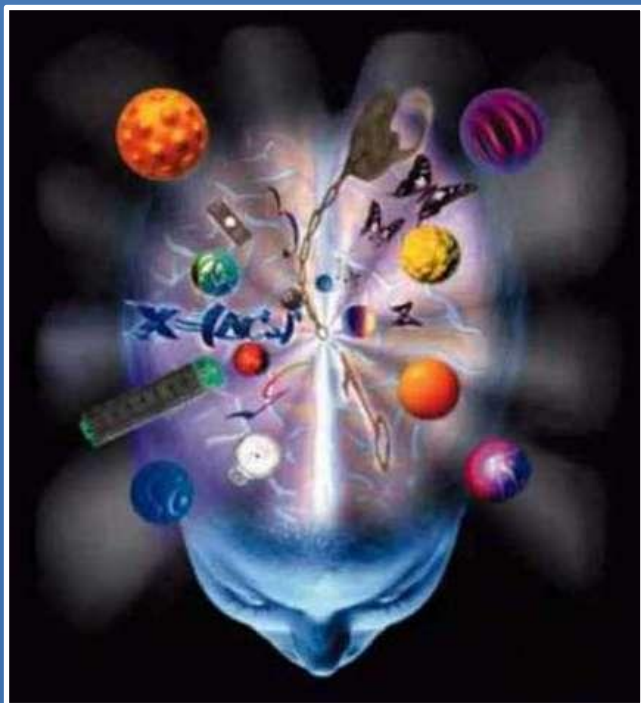
Не для всех людей естественное взаимодействие является одинаковым



Распознавание образов Общие сведения

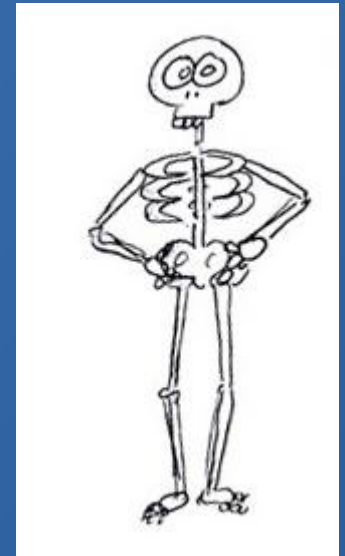
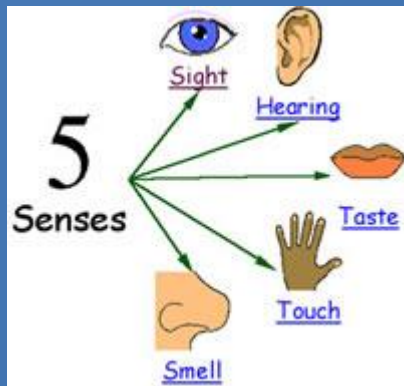


Распознавание образов (объектов, сигналов, ситуаций, явлений или процессов) - это самая распространенная задача, которую человеку приходится решать практически ежесекундно от первого до последнего дня своего существования.



Принятие решений

В повседневной деятельности человек постоянно сталкивается с задачами, связанными с принятием решений, обусловленных непрерывно меняющейся окружающей обстановкой.





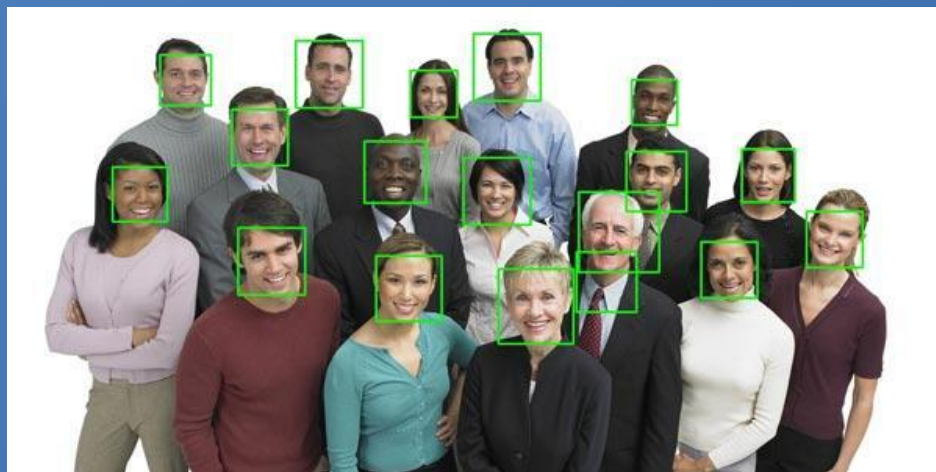
Применение распознавания образов

Большие возможности современной вычислительной техники позволяют в реальном масштабе времени обрабатывать огромное количество информации.

- устройства и машинные программы для чтения типографских и машинописных букв
- классификация электрокардиограмм
- распознавание слов из конечного словаря, произнесенных одним и тем же диктором
- выделение в речи ключевого слова, произнесенного разными дикторами



Применение распознавания образов (2)



Применение распознавания образов

Распознавание образов (объектов, сигналов, ситуаций, явлений или процессов) – это задача идентификации объекта или определения каких-либо его свойств по его изображению (оптическое распознавание) или аудиозаписи (акустическое распознавание) и другим характеристикам.

Образ - это классификационная группировка в системе классификации, объединяющая (выделяющая) определенную группу объектов по некоторому признаку.

Образы обладают характерным свойством, проявляющимся в том, что ознакомление с конечным числом явлений из одного и того же множества дает возможность узнавать сколь угодно большое число его представителей



Методы распознавания образов

- **МЕТОДЫ РАСПОЗНАВАНИЯ ОБРАЗОВ**
 - **Метод перебора**
 - **Метод анализа характеристик**
 - **Использование искусственных нейронных сетей**



Метод перебора

Производится сравнение с базой данных, где для каждого вида объектов представлены всевозможные модификации отображения.

Примеры,

- для оптического распознавания образов можно применить метод перебора вида объекта под различными углами, масштабами, смещениями, деформациями и т. д.;
- для букв нужно перебирать шрифт, свойства шрифта и т. д.;
- для распознавания звуковых образов, соответственно, происходит сравнение с некоторыми известными шаблонами (например, слово



Метод анализа характеристик образа

При втором методе производится более глубокий анализ характеристик образа.

Примеры,

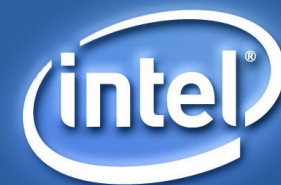
- в случае оптического распознавания это может быть определение различных геометрических характеристик.
- Звуковой образец в этом случае подвергается частотному, амплитудному анализу и т. д.

Метод не пригоден при изменении самого объекта или при низком качестве изображения.



Использование искусственных нейронных сетей

- Нейронная сеть представляет собой машину, моделирующую способ обработки мозгом конкретной задачи
- В основе идеи использования искусственных нейронных сетей лежит понятие пластичности головного мозга
- Искусственные нейронные сети используют множество взаимосвязей между элементарными ячейками вычислений – нейронами.



Использование искусственных нейронных сетей (2)

Искусственная нейронная сеть – это громадный распределенный процессор, состоящий из элементарных единиц обработки информации, накапливающих экспериментальные знания и представляющих их для последующей обработки.

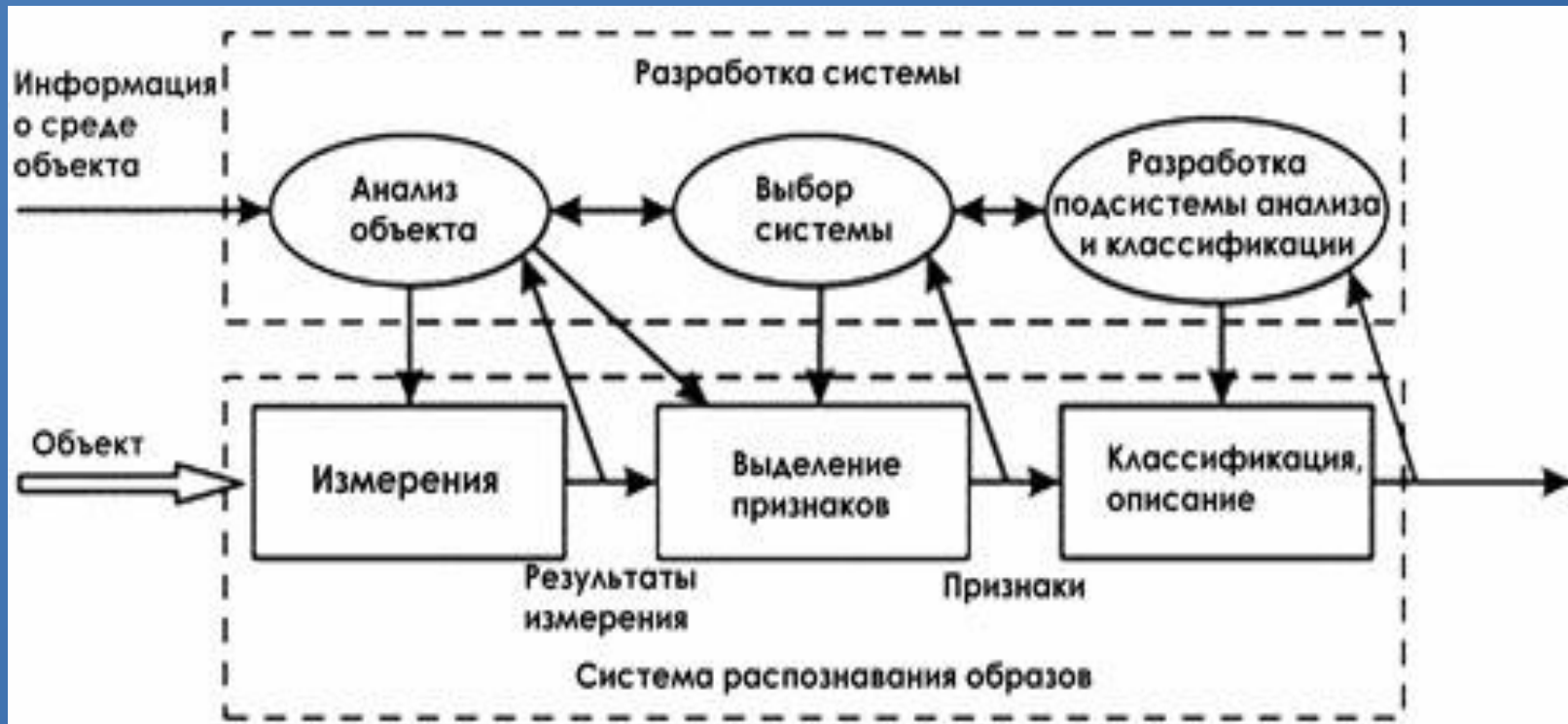


Использование искусственных нейронных сетей (2)

- Процедура, используемая для процесса обучения, называется алгоритмом обучения.
- Процедура обучения выстраивает в определенном порядке синаптические веса нейронной сети для обеспечения необходимой структуры взаимосвязей нейронов.
- Изменение синаптических весов представляет собой традиционный метод настройки нейронных сетей.
- Использование метода искусственных нейронных сетей требует либо большого количества примеров задачи распознавания при обучении, либо специальной структуры нейронной сети, учитывающей специфику данной задачи.
- Метод ИИНС отличается более высокой эффективностью и производительностью.



Общая структура системы распознавания



Этапы задач распознавания образов

- 1) преобразование исходных данных к виду, удобному для распознавания;
- 2) собственно распознавание (указание принадлежности объекта определенному классу).



Задачи распознавания образов

- **Распознавание** - отнесение предъявленного объекта по его описанию к одному из заданных классов;
- **Автоматическая классификация** - разбиение множества объектов, ситуаций, явлений по их описаниям на систему непересекающихся классов (таксономия, кластерный анализ, самообучение);
- **Выбор информативного набора признаков при распознавании;**
- **Прогнозирование**



ОСНОВЫ распознавания жестов



Жест

- Один из компонентов человеческой речи;
- Придает речи особую эмоциональность;
- Для слышащих людей жест – это дополнительное средство выразительности речи;
- Для людей, лишенных слуха, - основное средство общения





Отслеживание рук с помощью оборудованной перчатки

- Оборудованные перчатки, основанные на экзоскелетах, и оснащены более чем десятком датчиков (например, акселерометром, сенсорами, передающими тактильные, температурные и вибрационные ощущения, датчиками давления и др.).
- Информация о движении рук считывается с датчиков и передается на обработку в ПК.
- Полученная ПК информация анализируется и преобразовывается в символы.
- Данные о характеристиках отдельных жестов используется системой в качестве шаблонов.



Отслеживание рук с помощью оборудованной перчатки (2)



P5 Data Glove



Immersion Cyberglove



Отслеживание рук с помощью оборудованной перчатки (3)



Комплекс Fingual (Рисунок 5) японских исследователей из Университета Осаки и Университета Шиншу



Отслеживание рук с помощью оборудованной перчатки (4)

Преимущества

- Обеспечивают высокую точность отслеживания рук и распознавания жестов

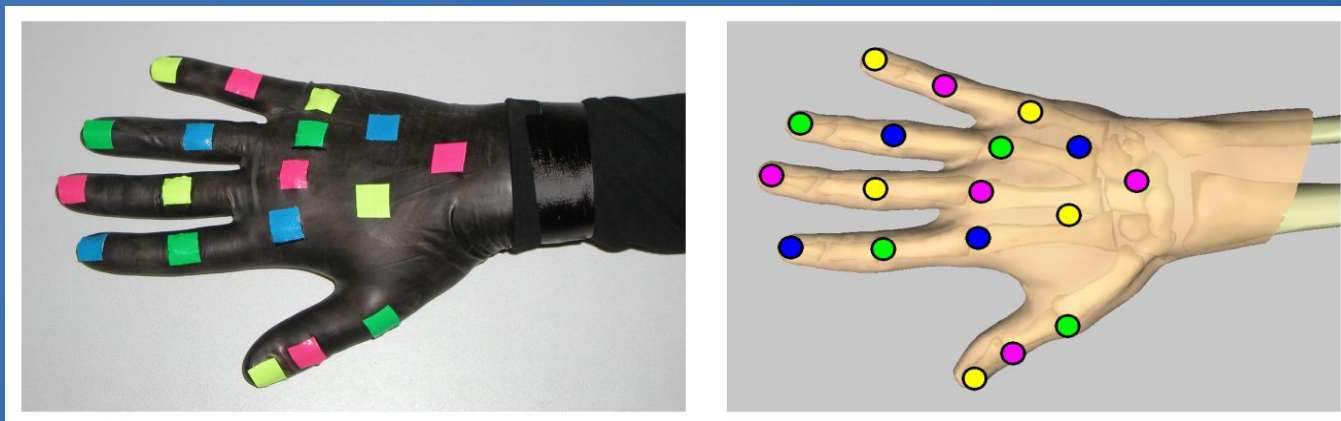
Недостатки

- Системы подобного рода являются дорогостоящими
- Применяющиеся перчатки очень громоздки, что ограничивает в движении.

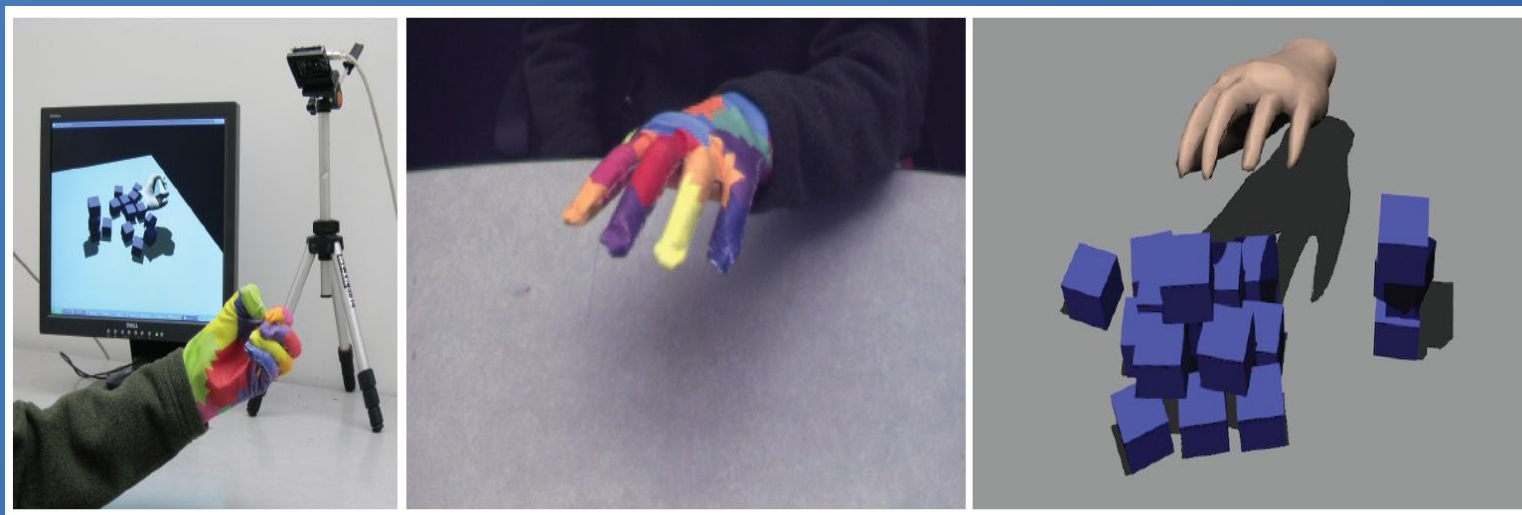


Захват движения на основе маркеров

- В основе лежит специальное оборудование;
- На руку человека прикрепляются датчики;
- Данные движения рук с датчиков фиксируются камерами и поступают на ПК, где сводятся в единую трёхмерную модель, точно воспроизводящую жесты человека, позже (или в режиме реального времени) на основе модели создаётся анимация жестов человека.



Захват движения на основе маркеров (2)



Демонстрирование работы системы, которая реконструирует руку на основе руки одетой в маркированную перчатку

Захват движения на основе маркеров (3)

Преимущества

- Подобные системы ориентированы на точность за счет простоты использования и установки.

Недостатки

- Использование навязчивых светоотражающих или цветных маркеров, или светодиодов и дорогих многокамерных установок.



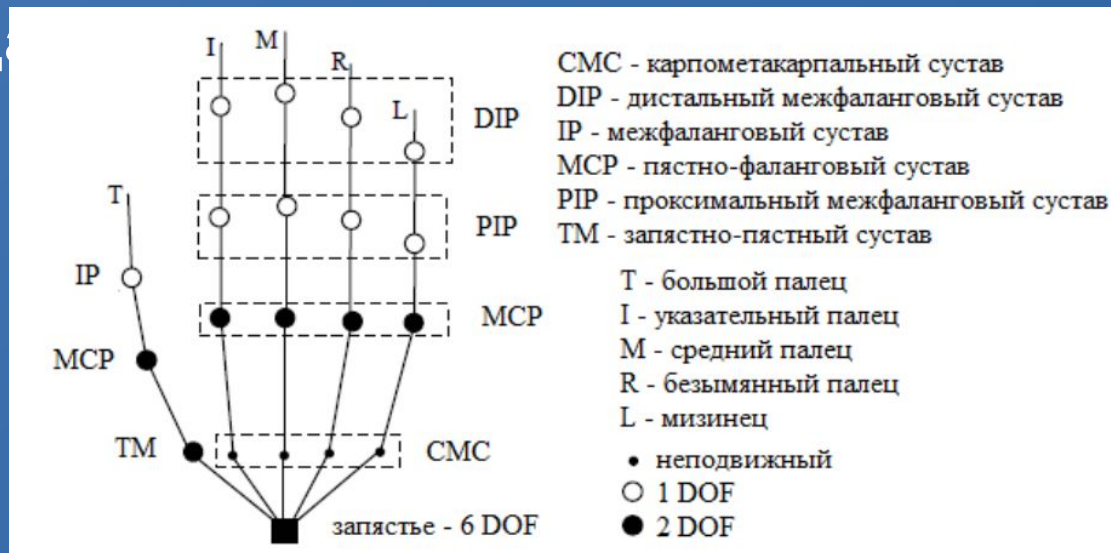
Визуальные методы распознавания жестов

- Визуальные методы распознавания жестов являются наиболее активной и перспективной областью исследований.
- Отслеживание голых рук основано на методах, работающих исключительно с визуальной информацией и не предъявляющих специальных требований к оснащению пользователя дополнительным оборудованием.
- Выделения контуров и силуэтов являются наиболее распространенными характеристиками, используемыми для идентификации позы руки.



Модели кисти с 27 степенями свободы

- Рука представляется как соединение составляющих частей, и каждое состояние называется «степенью свободы».
- Степени свободы определяют позиции и ориентации реального объекта.
- Используется упрощенная модель руки с 27 степенями свободы, ставшая неофициальным стандартом.



Модели кисти с 27 степенями свободы

- Системы отслеживания руки в пространстве на основе 3D модели руки используют заранее созданные базы данных всех известных конфигураций руки и их визуализации.
- Каждая конфигурация содержит множество углов наклона суставов и параметры позы руки для описания состояния руки.

Преимущества:

- Методы основанные на модели руки с 27 степенями свободы являются наиболее перспективными, так как они подразумевают полное оценивание позы и динамики руки.

Недостатки:

- Основными ограничениями подобных методов являются большая вычислительная сложность и ограниченная точность восстановления модели руки из-за наличия окклюзий, что делает невозможным их применение на



Статистические методы

- Предлагают построение признакового описания входного изображения и дальнейшую классификацию жестов именно по этому описанию.
- Решающую роль в методах подобного рода играет выбор подходящих характеристических признаков.

Например, для статических поз руки, можно использовать геометрические признаки, такие как кончики пальцев, направление пальцев, контур руки, а также негеометрические признаки (цвет кожи, форма, текстура и др.).

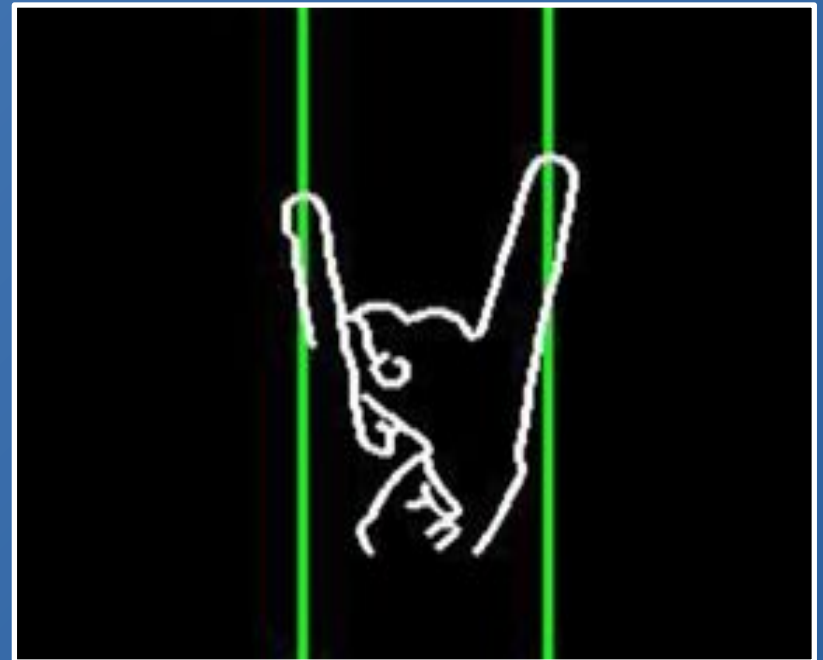
- Однако эти признаки не всегда доступны или адекватны для распознавания из-за взаимного перекрытия элементов руки и/или из-за условий освещения.



Статистические методы



После обнаружения
краев



Определенные пики
пальцев

Метрические методы

- Предполагают построение некоторой метрики на множестве входных изображений и выполнение классификации за счет сравнения входного изображения с набором эталонов.

Например, предлагается метрика, характеризующая степень сходства скелетов силуэтов ладони, и выполняется классификация жестов с помощью метода ближайшего соседа.



Примеры сценариев

- управление развлекательными приложениями и системами;
- управление медицинским оборудованием при требованиях стерильности, которые можно удовлетворить за счет минимизации тактильного контакта между оператором и устройством;
- визуализация и моделирование сложных трехмерных данных, удобство работы с которыми ограничено при применении стандартных двухмерных манипуляторов, таких как компьютерная мышь
- компьютерное распознавание языка глухонемых



Примеры сценариев (2)



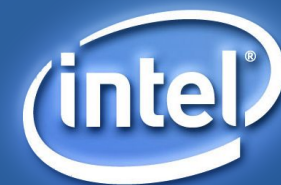
Примеры сценариев (3)



**Intel Perceptual Computing SDK
расширения границ
взаимодействия людей с
ограниченными возможностями с
внешним миром**



Intel® Perceptual Computing SDK 2013



Модуль распознавания жестов (1)

Входные потоки:



Модуль распознавания жестов (2)

4 типа обработки данных:

- BLOB
- Геометрические узлы
- Распознавание жестов
- Информационные сообщения



Модуль распознавания жестов (3)

VLOB информация представляет результаты промежуточной обработки изображений на входе. Типичные VLOB данные – обработанное глубинное изображение поля зрения камеры.

Геометрические узлы - это каркасные соединения (скелетон) человеческого тела или отдельные его части. SDK возвращает позиции следования геометрических узлов и другую полезную информацию.

Модуль SDK распознает набор predeterminedных поз и жестов, и возвращает результаты этого распознавания.

Поза - это устойчивое, долговременное положение тела

Жесты - это движения между позами.

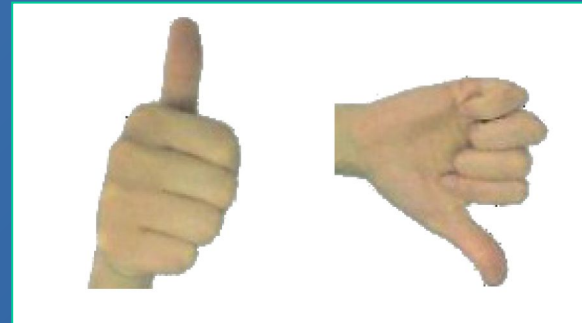
SDK модуль отправляет приложению оповещения об обнаружении ошибок, распознании объекта



Основные распознаваемые позы (1)

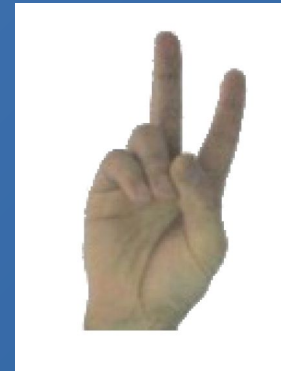
Открытость - с помощью просмотра атрибутов LABEL_OPEN и LABEL_CLOSE можно распознать, открыта или закрыта ладонь.

«Большой палец вверх» или «Большой палец вниз» - атрибуты LABEL_POSE_THUMB_UP и LABEL_POSE_THUMB_DOWN



Основные распознаваемые позы (2)

Знак победы - атрибут LABEL_POSE_PEACE.

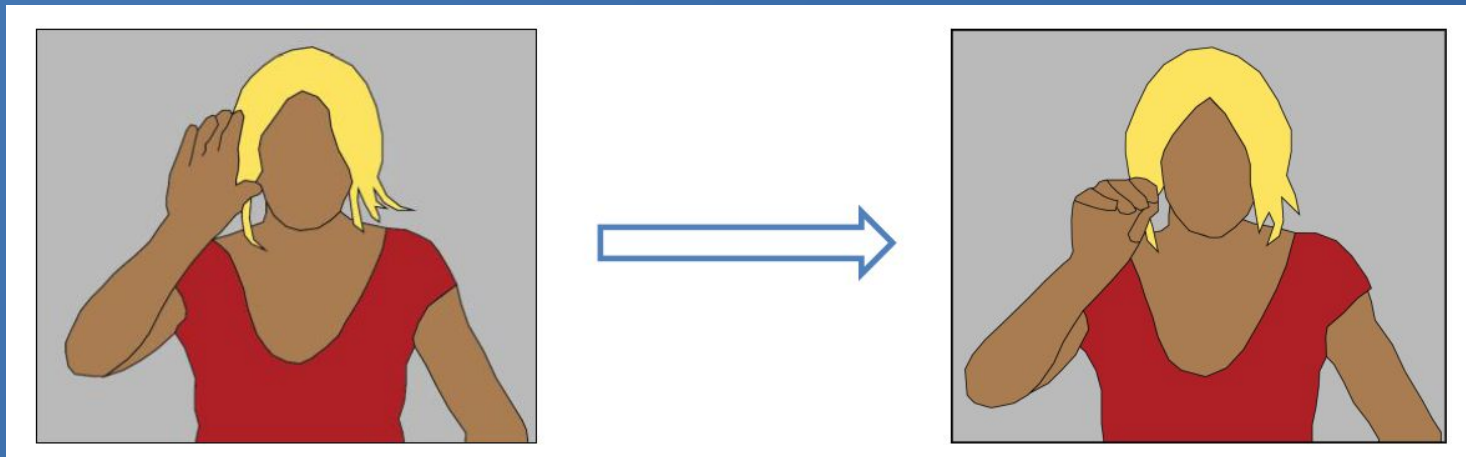


Большая пятерка – атрибут LABEL_POSE_BIG5



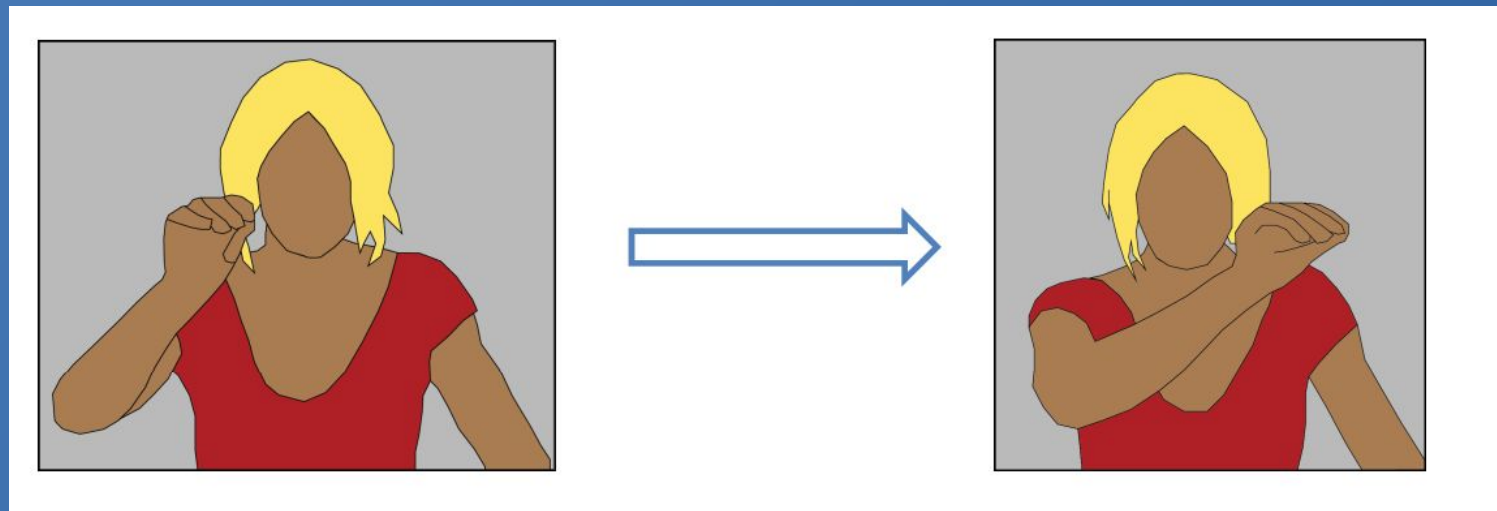
Основные распознаваемые жесты (1)

«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.



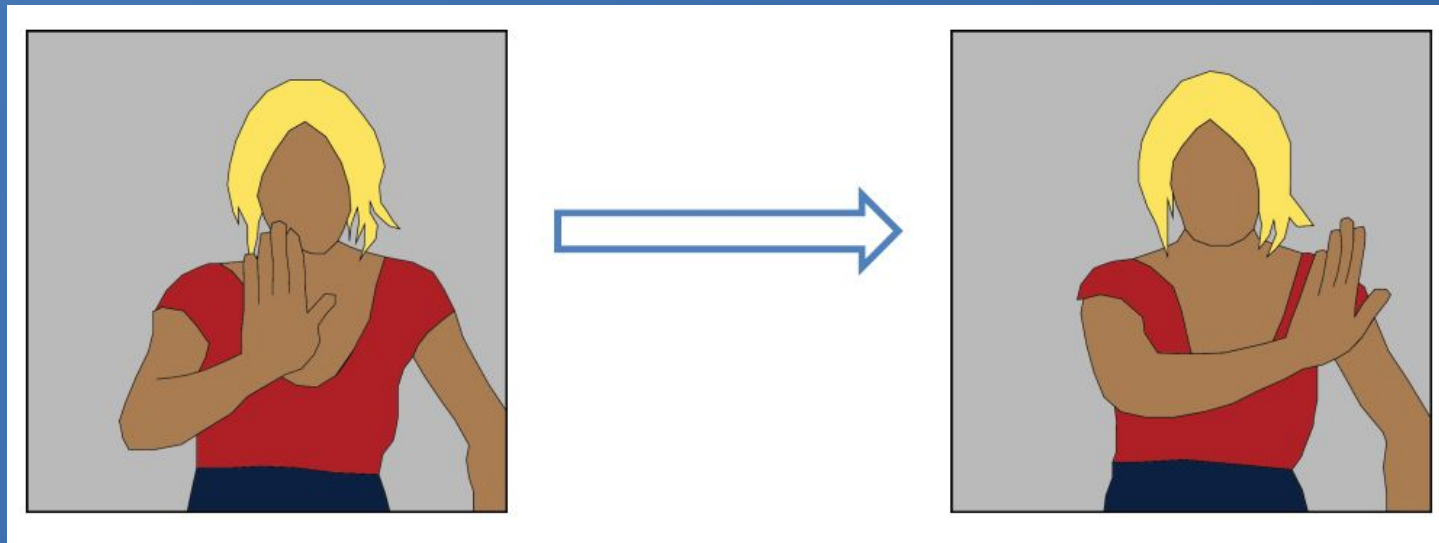
Основные распознаваемые жесты (2)

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



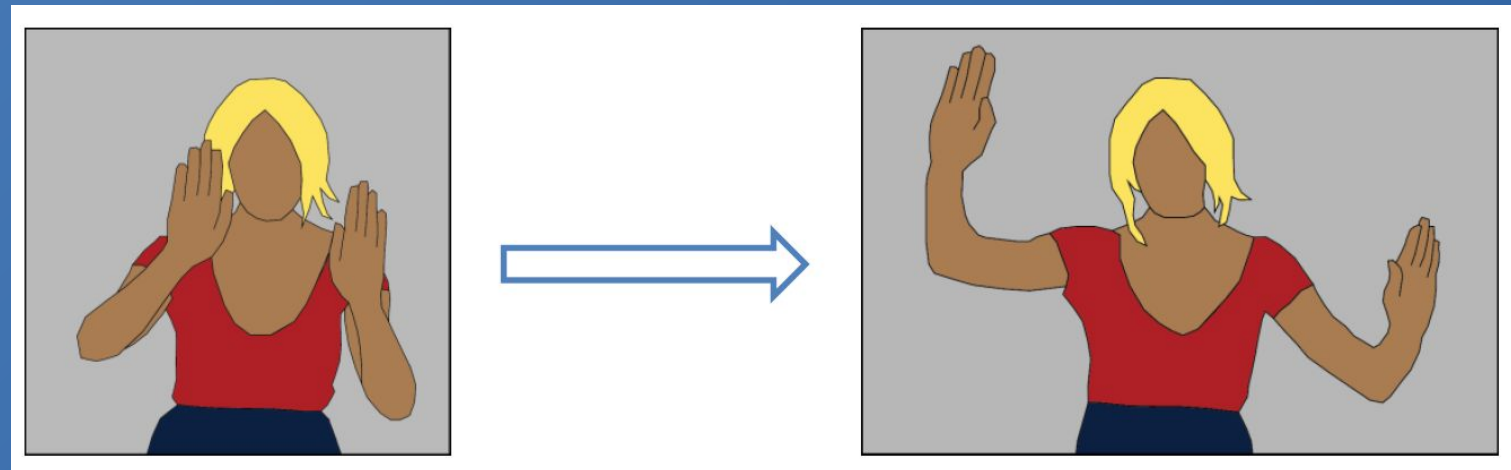
Основные распознаваемые жесты (3)

«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.



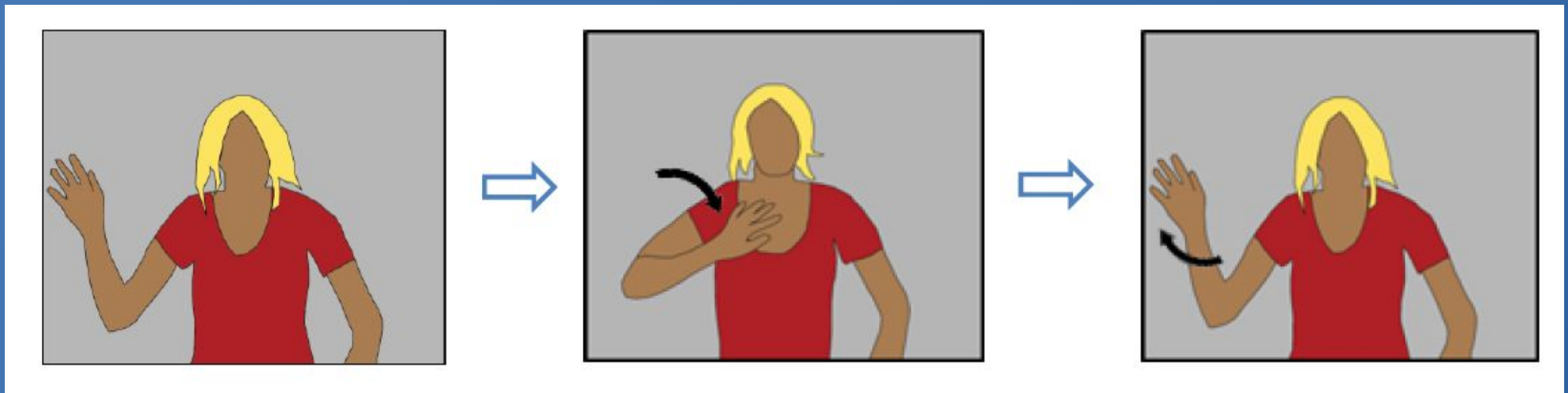
Основные распознаваемые жесты (5)

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



Основные распознаваемые жесты (6)

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



Основные распознаваемые жесты (7)

«Скольжение» - Скольжение является основным жестом навигации. Различают скольжение как слева направо, так и справа налево. Также выделяют скольжение снизу вверх и сверху вниз.

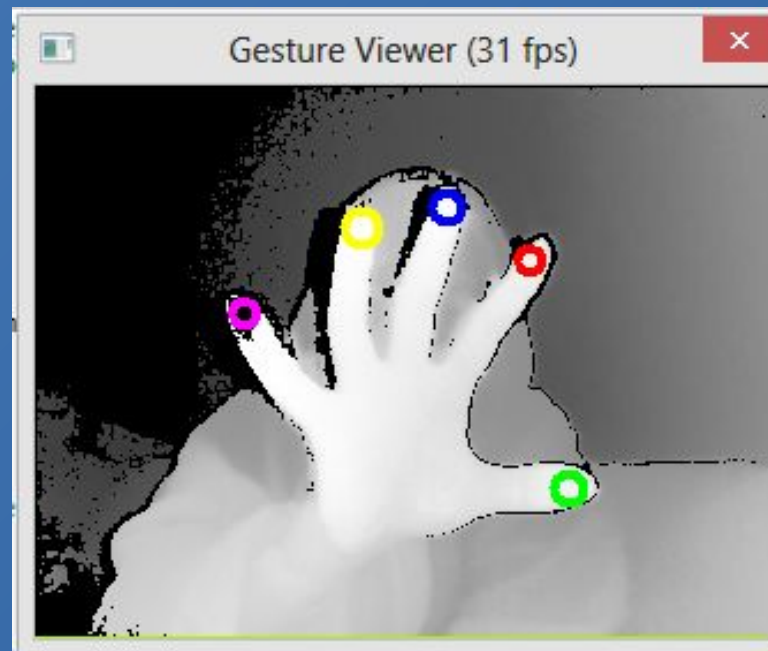
В SDK их можно найти, как LABEL_NAV_SWIPE_LEFT, LABEL_NAV_SWIPE_RIGHT, LABEL_NAV_SWIPE_UP и LABEL_NAV_SWIPE_DOWN, соответственно.



Отслеживание геометрических узлов (1)

SDK позволяет определить

- точки центра ладони;
- кончиков пальцев;
- самую дальнюю точку



Отслеживание геометрических узлов (2)

Для получения координат каждого пальца используется функция QueryNodeData

Пример 1: получение координат x и y указательного

```
PXCGesture::GeoNode indexNode;  
Gest->QueryNodeData(0,  
PXCGesture::GeoNode::LABEL_BODY_HAND_PRIMARY |  
PXCGesture::GeoNode::LABEL_FINGER_INDEX, &indexNode);  
wprintf(L"index x=%f\n", indexNode.positionImage.x);  
wprintf(L"index y=%f\n", indexNode.positionImage.y);
```



Отслеживание геометрических узлов

(3)

Детализирующие метки для получения координат других пальцев и центра ладони:

PXCGesture::GeoNode::LABEL_FINGER_THUMB – большой палец
PXCGesture::GeoNode::LABEL_FINGER_INDEX – указательный палец
PXCGesture::GeoNode::LABEL_FINGER_MIDDLE – средний палец
PXCGesture::GeoNode::LABEL_FINGER_RING – безымянный палец
PXCGesture::GeoNode::LABEL_FINGER_PINKY – мизинец

PXCGesture::GeoNode::LABEL_HAND_MIDDLE – центр ладони



Отслеживание геометрических узлов (4)

Пример 2: получение координат x и y центра ладони

```
PXCGesture::GeoNode centr; //центр ладони
Gest->QueryNodeData(0, PXCGesture::GeoNode::LABEL_BODY_HAND_PRIMARY |
PXCGesture::GeoNode::LABEL_HAND_MIDDLE, &centr);
wprintf(L"centr x=%f\n", centr.positionImage.x);
wprintf(L"centr y=%f\n", centr.positionImage.y);
```

Пример 3: получение координат x и y центра ладони

```
PXCGesture::GeoNode open;
gest->QueryNodeData (0,PXCGesture::GeoNode::LABEL_BODY_HAND_PRIMARY,
&open);
wprintf_s(L"Openness: %s (%d%%)\n",
open.opennessState==PXCGesture::GeoNode::LABEL_OPEN?L"Open":
(open.opennessState==PXCGesture::GeoNode::LABEL_CLOSE?L"CLOSE":L"Unknow
n"), open.openness);
```



Контрольные вопросы

1. Что означает термин «информационная доступность»?
2. В каких областях применяется распознавание образов?
3. Какие существуют методы распознавания образов?
4. Назовите основные элементы системы распознавания образов?
5. Какие существуют методы распознавания жестов?
6. В чем заключаются преимущества и недостатки каждого метода?
7. Приведите примеры систем, основанных на распознавании и отслеживании жестов.
8. Какие позы уже заложены в функционале Intel Perceptual Computing SDK?

