

Обнаружение пальцев руки в дальностных изображениях

Аннотация. Рассмотрена технология распознавания пальцев руки человека в видеоряде, на основе которой предлагается строить системы взаимодействия человека с ЭВМ в реальном времени.

Ключевые слова: распознавания пальцев, дальностные изображения.

Введение

В течение жизни человек получает и передает информацию. Язык общения человека и компьютера определялся в основном современным техническим уровнем и ограничен технологическими возможностями устройств ввода-вывода информации, к которым приходится приспосабливаться [1-5]. С развитием информационных технологий язык общения также прогрессирует, становится более дружественным к человеку за счет совершенствования и адаптации программно-аппаратных средств. Это может быть, например, общение, построенное на основе автоматического распознавания речевых сигналов или жестов человека. В настоящей работе рассматривается подход к созданию системы взаимодействия с ЭВМ на основе понимания жестов. В частности, описывается алгоритм, который позволяет обнаружить пальцы руки в дальностных изображениях и видеорядах [6]. Алгоритм работает как для открытой ладони, так и для случаев, показа отдельных пальцев и их комбинаций.

1. Дальностные изображения

Дальностные изображения (или карты глубин), в отличие от обыкновенных изображений, содержат информацию не о яркости и/или цвете, а о расстоянии (глубине) от каждого пикселя изображения до камеры наблюдения [6]. У камер,

позволяющих получить дальностные изображения, есть одно значимое превосходство: по сравнению с обыкновенными камерами дальностные камеры почти не чувствительны к изменениям освещения наблюдаемой среды [7, 8].

В зависимости от типа сенсора может меняться структура пикселя дальностного изображения. Например, для реализации алгоритма, предложенного в настоящей работе, был использован сенсор Kinect компании Microsoft [9]. Сенсор предназначен для использования во внутренних помещениях, причем максимальное расстояние, которое он может просматривать составляет 1000 см [7], а каждый пиксель дальностного изображения, возвращаемый Kinect-ом, хранится в двух байтах.

Существуют преобразования, которые позволяют, на основе дальностного изображения, получить цветное или полутоновое изображение, которое не зависит от структуры пикселя. Пусть имеется дальностное изображение *DepthMap*, размер которого равен *XRes*×*YRes*. Построим новое изображение –*GrayScale*, у которого значения пикселей определяются следующим образом:

$$GrayScale [row, column] = 255 * \frac{DepthMap [row, column]}{MaxDist},$$

где *MaxDist* - максимальное расстояние пикселя от сенсора, $0 \leq row \leq YRes$, $0 \leq column \leq XRes$.

Из уравнения следует, что пиксели нового изображения имеют значения в диапазоне

[0,255]. Результатом преобразования является полутоновое изображение (8btt - bits per pixel) (Рис. 1).

Предложенный метод преобразования, разумеется, не является единственным. На Рис. 2 показано изображение человека, построенное с помощью гистограммы глубины и сглаживания. В дальнейшем, дальностные изображения будут отображаться в виде монохромных изображений с помощью одного из указанных методов.

Рассмотрим несколько свойств, полученных монохромных изображений:

1. Изображения (Рис. 1, Рис.2) будут одинаковыми, как при дневном свете, так и при полной темноте. Это означает, что алгоритм распознавания пальцев не будет зависеть от освещения окружающей среды.

2. На полученных изображениях видна некая тень. Причиной ее появления является не источник освещения, а особенности камеры, с помощью которой получаются исходные дальностные изображения (Рис. 3).

Почти во всех устройствах получения дальностных изображений используется оптимизированная по затратам схема с одним проектором, который излучает инфракрасный свет, и сенсором, который получает отраженный со сцены свет. Это вносит свои ограничения, так при наличии перед камерой объектов, которые заслоняют друг друга, возникает некая полоса, которая не отражает свет проектора сенсору (Рис. 3) [10].

2. Задача распознавания пальцев руки

Рассмотрим далее задачу распознавания пальцев руки человека. Алгоритм ее решения должен, получая на вход дальностное изображение (видеоряд), выдавать трехмерные координаты всех кончиков пальцев руки. Видеоряд представляет собой список изображений, привязанный ко времени. На сегодняшний день существуют корректно работающие алгоритмы и программно-технические средства, которые распознают части тела человека (Рис. 4, Рис.5). К их числу относятся, например, OpenNI [11], XBox [12], Microsoft Kinect SDK [13]. Можно предположить, что наличие данных средств является достаточным для определения местонахождения ладони человека, что облегчает решение поставленной задачи.

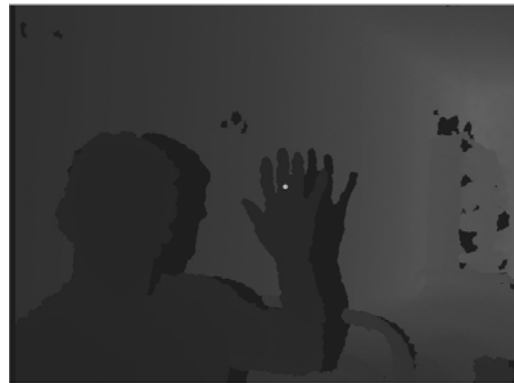


Рис. 1. Полутоновое изображение человека, полученное на основе формулы



Рис. 2. Полутоновое изображение человека, полученное с помощью гистограммы глубины и сглаживания

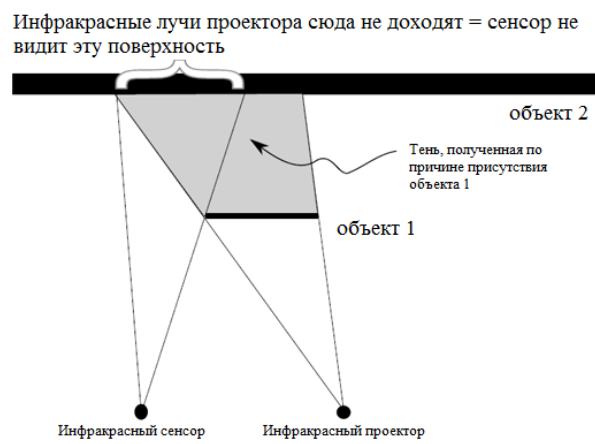


Рис. 3. «Тень», полученная в дальностном изображении

3. Описание алгоритма решения задачи

К алгоритму распознавания пальцев руки на изображениях предъявляются следующие требования:

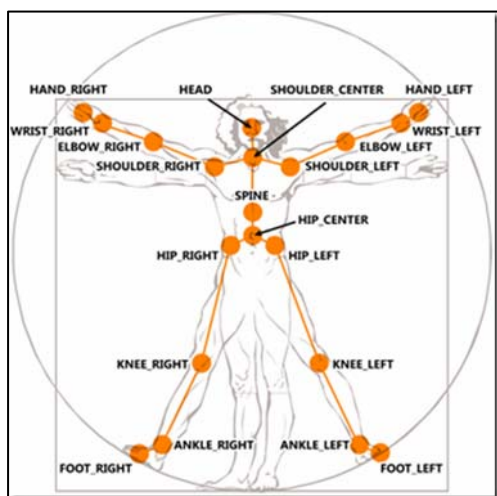


Рис. 4. Части тел, распознанные с помощью MS Kinect SDK [13]

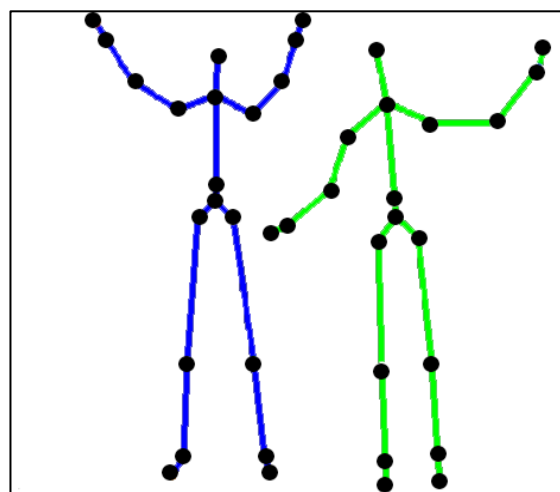


Рис. 5. Скелеты, построенные на основе данных полученных от MS Kinect SDK [13]

- скорость обработки должна обеспечивать распознавание пальцев руки в видеорядах, что соответствует времени обработки кадра не более чем за 1/24 секунду;

- работа алгоритма не должна зависеть от точного местоположения точки ладони руки, передаваемой на вход алгоритма;

- алгоритм должен возвращать координаты пальцев руки в трехмерном пространстве.

Предлагаемый алгоритм состоит из нескольких этапов, показанных на Рис.6. Для определения расстояния между двумя точками используется метрика Евклида в трехмерном пространстве.

4. Обработка входного дальностного изображения

Пусть на вход алгоритма подается дальностное изображение. Если известны координаты некоторой точки руки, то можно построить новое изображение, в котором будут отсутствовать многие несущественные данные. Ладонь человека имеет конечные размеры. Это означает, что существует верхняя граница R для расстояния между любыми двумя точками ладони. Отсюда следует, что если построить сферу, приняв за центр известную точку руки, с радиусом равным R , то все точки ладони будут

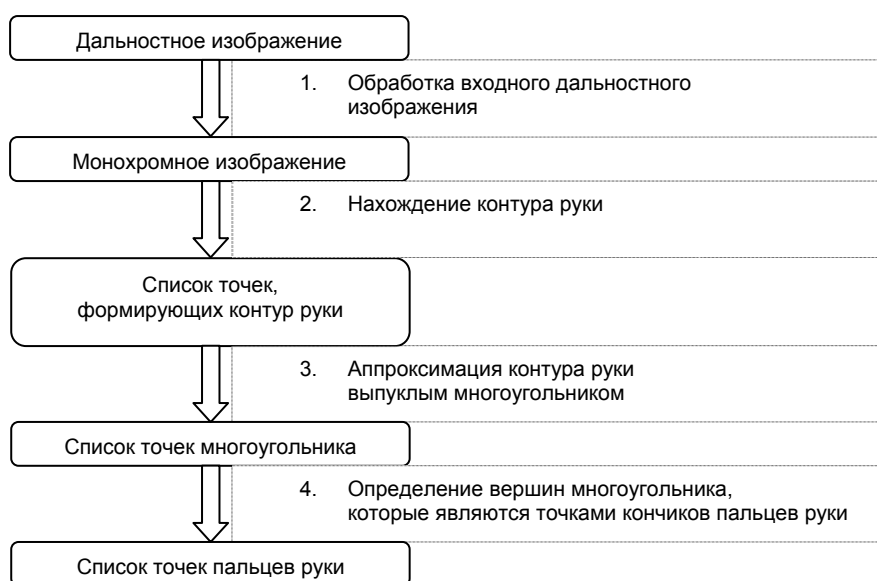


Рис.6. Схема алгоритма распознавания

находиться в этой сфере (в том числе точки, принадлежавшие кончикам пальцев).

Получая входное изображение, построим новое, где пикселям, не находящимся внутри сферы, будут присвоены нулевые значения. На Рис. 7 показан результат такой конвертации. Можно заметить, что в полученном изображении отсутствуют «тень» и фон, которые наблюдались в конвертированных полутонных изображениях.

5. Нахождение контура руки

Для нахождения контура руки используется алгоритм оконтуривания изображения, предложенный Satoshi Suzuki [15], после чего из всех полученных контуров выбирается тот, в котором находится точка ладони. Результат показан на Рис. 8:

6. Аппроксимация контура руки выпуклым многоугольником

Алгоритм на этом шаге, получая на вход контур руки человека, аппроксимирует его выпуклым многоугольником (Рис.9). Для этого

используется механизм, предложенный Jack Sklansky [14].

7. Определение кончиков пальцев руки

Для определения кончиков пальцев строится многоугольник руки. Затем создаются круги с радиусом, линейно зависящим от расстояния руки, и центрами в вершинах многоугольника, которые не находятся на границе отрезка руки. Для каждого круга определяются процентные соотношения области, заполненной изображением руки и пустой области. На основе этого соотношения принимается решение о принадлежности точки кончику пальца (Рис. 10,а).

На Рис.11 видны круги с центрами в вершинах многоугольника. В жирных кружках процентное содержание области руки было настолько мало, что алгоритм определил их центры, как точки пальцев. В светлых кружках – не прошедшие проверку, по причине процентного соотношения руки с пустой областью; белые точки – не прошедшие проверку, по причине нахождения вершины на границе отрезка руки. Заметим, что алгоритм может найти несколько точек для одного и того же пальца.



Рис. 7. Обработка входного изображения

(a) Монохромное изображение в начальном дальностном изображении
(b) Конвертированное изображение



Рис. 8. Контур руки



Рис. 9. Выпуклый многоугольник руки

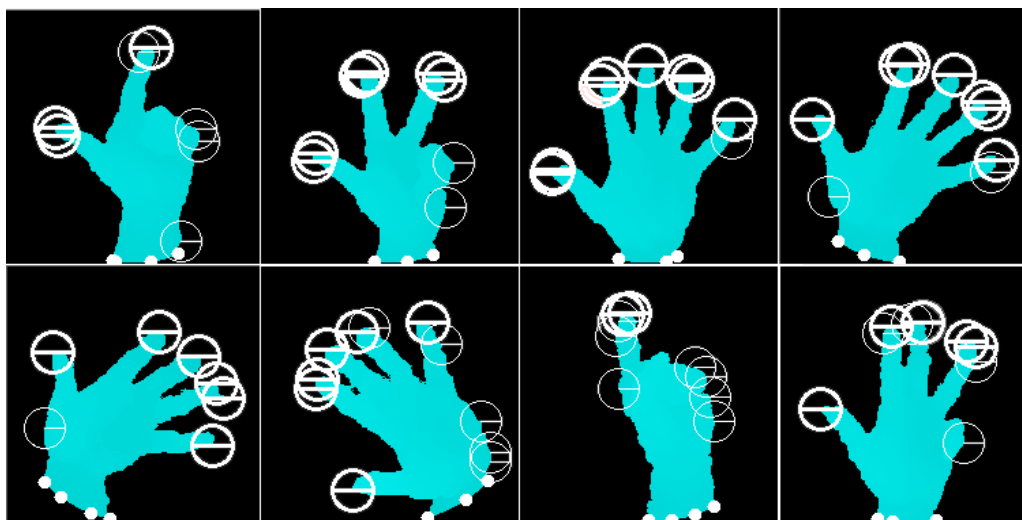


Рис.10.Круги с центрами в вершинах многоугольника

Таким образом, получаем двумерные координаты кончиков пальцев человека, как это показано на Рис. 11. Имея дальностное изображение и двумерные координаты пальцев руки можно найти и недостающие третьи координаты.

8. Тестирование алгоритма и его результаты

Предложенный алгоритм был реализован и протестирован на видеорядах. В таблице показаны результаты работы алгоритма для дальностного изображения с размером 640×480 на процессоре Intel T6500.

Тестирование показало, что алгоритм может применяться для видеорядов с 50 FPS (frames per second/число кадров в секунду). Этого достаточно для построения приложения, обеспечивающего общение человека с машиной в режиме реального времени.

Этап алгоритма	Время выполнения, мс
Обработка входного дальностного изображения	14
Нахождение контура руки	1
Аппроксимация контура руки выпуклым многоугольником	1
Определение вершин многоугольника, которые являются точками кончиков пальцев руки	4

Большая часть алгоритмов, рассмотренных в настоящей работе, пригодна для обработки обычных изображений с небольшими модификациями, что расширяет возможности практического применения предложенной технологии.

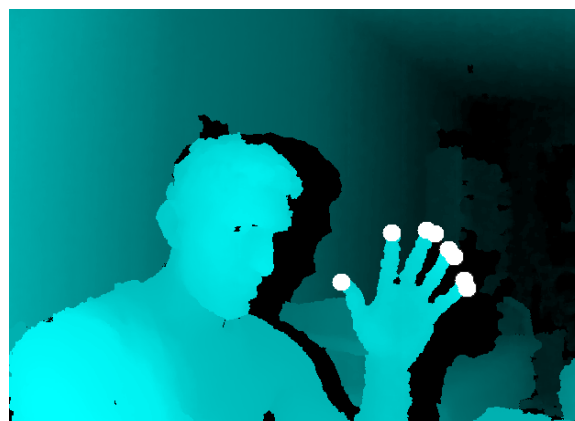


Рис. 11. Кончики пальцев руки

Литература

1. Alessandro Valii.Natural Interaction White Paper, p.3. <http://www.naturalinteraction.org/images/whitepaper.pdf>
2. United Stated Patent N 3,541,541. X-Y Position indicator for a display system. Nov 17, 1970 <http://www.google.com/patents?vid=3541541>
3. United States Patent N196.652 .Computer Keyboard. Oct. 22. 1963. <http://www.google.com/patents/about?id=OPFyAAAAEBAJ>
4. United States Patent N2002/0044137.SENSOR PANEL POSITION COORDINATES SIGNAL DETECTING CIRCUIT. <http://www.google.com/patents/about?id=vcuNAAAAEBAJ&dq=sensor+panel>
5. United States Patent N6.741.237 .Touch Screen. <http://www.google.com/patents/about?id=Ge0RAAAAEBAJ>
6. Форсайт. Понс.Компьютерное зрение.р. 637-640.
7. PrimeSense design specifications: <http://www.primesense.com/?p=514>
8. SR 4000 Data Sheet : <http://www.mesa-imaging.ch/prodview4k.php>
9. Microsoft Kinect for Xbox 360. <http://www.xbox.com/en-US/kinect>

10. OpenKinect FAQ - <http://openkinect.org> «Why are there black shadows in the depth image?»
http://media.zero997.com/kinect_shadow.pdf
11. The OpenNI organization. <http://www.openni.org>
12. Microsoft Xbox. www.xbox.com
13. Microsoft Kinect Sdk. <http://research.microsoft.com/en-us/um/redmond/projects/kinectsdk/>
14. Jack Sklansky Finding the convex hull of a simple polygon university of California, Irvine, CA 92717, U.S.A.
15. Suzuki S., Abe K. Topological Structural Analysis of Digitized Binary Images by Border Following.

Нагапетян Ваагн Эдвардович. Аспирант Российского университета дружбы народов. Окончил факультет информационных технологий Ереванского государственного университета в 2010 году. Область научных интересов: искусственный интеллект, машинная графика, распознавания образов, машинное обучение. Email: vahagnahapetyan@gmail.com.