

УДК 004.925: 004.932: 004.942: 004.946

ИНТЕРФЕЙС БЕСКОНТАКТНОГО УПРАВЛЕНИЯ**Котюжанский Л.А.***ФГАОУ ВПО «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина», Екатеринбург, e-mail: matveeva-umc@yandex.ru*

Предлагаемый интерфейс реализует управление мультимедиа и любым программным обеспечением при помощи различных жестов, осуществляемых одним или несколькими пользователями. В качестве источника данных выступают сенсоры-дальномеры, такие, как Microsoft Kinect, Asus Xtion. Одной из самых главных задач разрабатываемого интерфейса является отслеживание рук, находящихся на некотором расстоянии от сенсора и распознавание пальцев. На основе полученных данных (координаты рук и пальцев) вычисляются скорости, траектории движения, поворот рук пользователя. Затем полученные данные ассоциируются с различными командами приложений. Так, например, две руки с пятью видимыми пальцами могут ассоциироваться с командой закрыть или открыть определенное приложение. В статье показаны основные моменты работы данного интерфейса, описаны ключевые вычисления, в заключении перечислены основные проблемы в работе системы и возможные пути их решения.

Ключевые слова: стереопара, сенсор Kinect, детектирование жестов, сегментирование, блов, шейдер GLSL**INTERFACE CONTACTLESS CONTROL****Kotyuzhanskiy L.A.***Ural Federal University named after the first President of Russia B.N. Yeltsin, Ekaterinburg, e-mail: matveeva-umc@yandex.ru*

The proposed interface is implemented and in any media control software with different gestures by one or more users. The source of data are the sensors, range finders, such as the Microsoft Kinect, Asus Xtion. One of the main tasks is to track the development of interfaces hands, some distance from the sensor and fingerprint recognition. Based on these data (coordinates of hands and fingers) are calculated speed, trajectory, turning the user's hands. Then, the data associated with different teams applications. For example, two hands with five fingers visible may be associated with a team of close or open a specific application. The article describes the main points of this interface describes the key calculation in the conclusion are the main problems in the system and their possible solutions.

Keywords: stereo pair, the sensor Kinect, gesture detection, segmentation, blob, shader GLSL

Естественный человеко-машинный интерфейс представляет собой пользовательский интерфейс, который является для пользователя объективно невидимым или становится невидимым в процессе обучения. Наиболее характерной особенностью естественных интерфейсов является отсутствие физических клавиатуры и мыши. Традиционные компьютерные интерфейсы опираются на использование искусственных устройств управления, для работы с которыми пользователю требуется время для обучения. Концепция естественных интерфейсов предполагает способность пользователя осуществлять естественные движения или жесты, что позволит практически сразу успешно управлять компьютерными приложениями или манипулировать содержимым экрана. Распознавание жестов имеет много преимуществ перед другими средствами ввода, такими, как клавиатура, мышь или сенсорный экран, в частности, большее количество степеней свободы для управления, отсутствие механического износа устройств управления, возможность одновременной работы нескольких пользователей.

Предложенная система распознавания движений может контролировать различные устройства и приложения с использовани-

ем набора простых жестов. Для некоторых типов устройств и приложений управление жестами является более практичным и интуитивно очевидным. Например, при управлении мобильным роботом основными командами типа «идти вперед (назад)», «увеличить (уменьшить) скорость»; при работе с графическим содержимым, включающей вращение, перемещение, масштабирование; в компьютерных играх и других приложениях, где происходит взаимодействие с виртуальной реальностью.

Данная работа предлагает достаточно детальное описание системы бесконтактного управления (рис. 1) при помощи пальцев и рук пользователей, их взаимного расположения и движения. Система отслеживает пальцы и ладони пользователей в пространстве, определяет количество видимых пальцев, угол поворота руки, анализирует различные жесты, формируемые определенными комбинациями пальцев на руках пользователей с учетом скорости и ускорения их движения. В качестве источника данных используется стереопара или сенсор Kinect [1].

Основная идея данной системы – формирование различных команд управления на основе определения координат, скорости, ускорения видимых системой пальцев и ла-

дней пользователя, идентификации левой и правой ладони. При этом формируемые команды передаются по локальной сети при помощи специально разработанного протокола передачи данных о жестах – GUIO (Gesture User Interface Output) от «распоз-

навателя» (программа детектирования жестов) до «приемника» (рис. 2).

Приложением «приемник» может являться любое прикладное программное обеспечение: операционная система, мультимедиа презентация, игра.

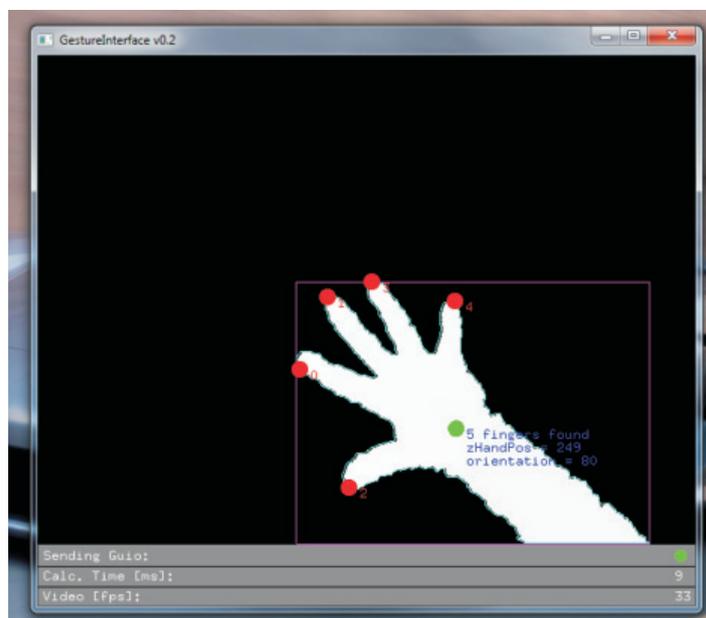


Рис. 1. Интерфейс бесконтактного управления программным обеспечением с помощью жестов



Рис. 2. Диаграмма взаимодействия «распознавателя» и прикладного программного обеспечения в интерфейсе бесконтактного управления

В процессе работы над системой были отобраны для распознавания несколько жестов, которые обрабатывались на специально написанном графическом приложении. Весь набор действий системы бесконтактного управления можно разделить на два больших блока – получение первичных данных (рис. 3) и получение вторич-

ных данных (рис. 4). При первичной обработке анализируются данные сенсора для выделения и отслеживания рук и пальцев пользователя. Вторичная обработка использует результаты первичной обработки. Вычисляются скорости пальцев и рук, определяются повороты рук, распознается жест – «клик», выполняется определение

количества пальцев. Эти данные накапливаются и отправляются в прикладное приложение по протоколу GUIO. Данный

протокол базируется на протоколе UDP [2] и позволяет отсылать информацию о жестах по локальной сети.



Рис. 3. Основные шаги для получения первичных данных – определение рук и пальцев пользователя



Рис. 4. Получение вторичных данных – количество пальцев, скорости рук и пальцев, угол поворота рук, удаленность руки и пальцев от сенсора, определение жеста «клик в воздухе»

Для получения системой первичных данных используются различные техники машинного зрения. В частности – сегментирование рук пользователя. Результатом

сегментирования является бинарное изображение белых рук на черном фоне. Получение вторичных данных отражено на рис. 5–6.



Рис. 5. Получение вторичных данных – количество пальцев, скорости рук и пальцев, угол поворота руки, удаленность руки и пальцев от сенсора, определение жеста «клик в воздухе»

Система бесконтактного управления была реализована с использованием различных программных библиотек. Так, в качестве каркаса приложения и вывода графических данных была выбрана библиотека Openframeworks, которая позволяет быстро разрабатывать любые интерактивные аудио-визуальные приложения. Часть алгоритмов была реализована на шейдерах GLSL. Для получения данных с сенсора Kinect использовалась открытая библиотека libfreenect [3]. Часть алгоритмов была реализована на OpenCV. Выбор этих библиотек обусловлен тем, что они являются открытыми, бесплатными и кроссплатформерными.

За счет широкого использования GLSL удалось добиться минимальной загрузки ЦПУ и времени обработки одной итерации. Так, загрузка ЦПУ Intel Core i3 в ОС Windows 7 составила порядка 15%, а время вычисления одной итерации 9 мс.

На текущий момент представленная система, протестированная на наборе практических задач [4–6], имеет ряд серьезных ограничений. В частности – требуется обеспечить четкое распознавание кистей пользователя, чтобы свести к нулю попытки системы найти пальцы на каких-либо других объектах кроме кисти или ладони пользователя. Также необходимо организовать формирование системы координат на основе биометрических данных пользователя. На-

пример, возможно создание некоей виртуальной прямоугольной координатной сетки, находящейся как бы перед лицом пользователя. Это делает удобным перемещение курсора по всему экрану путем незначительного перемещения ладони или пальца перед собой в пределах указанной сетки. Кроме того, желательно запоминать пользователя, чтобы исключить возможность срабатывания системы на случайных объектах или людях. Большая часть отмеченных проблем решается путем отслеживания ключевых, опорных точек пользователя: ладоней, плеч, локтей, головы. В ближайшее время планируется расширить недостающий функционал системы, используя результаты работы [7].

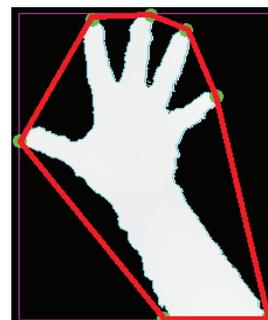
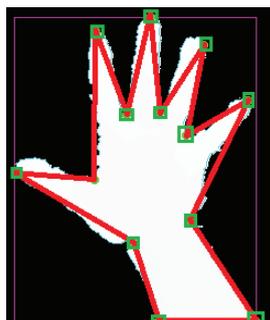
Список литературы

1. Xbox 360 Kinect Sensors and Bundles, [Электронный ресурс]. – Режим доступа: <http://www.microsoftstore.com> (дата обращения 14.01.2013).
2. UDP: User Datagram Protocol [Электронный ресурс]. – Режим доступа: <http://www.xserver.ru/computer/protocol> (дата обращения 14.01.2013).
3. Drivers and libraries for the Xbox Kinect device on Windows [Электронный ресурс]. – Режим доступа: <https://github.com/OpenKinect/libfreenect> (дата обращения 14.01.2013).
4. OceanGestureDemo [Электронный ресурс]. – Режим доступа: http://www.youtube.com/watch?v=gn_dnYbi4oE (дата обращения 11.03.2012).
5. JaguarCarGestureDemo [Электронный ресурс]. – Режим доступа: <http://www.youtube.com/watch?v=Utq6ZJFSzk> (дата обращения 15.03.2012).
6. GestureMovieControl [Электронный ресурс]. – Режим доступа: http://www.youtube.com/watch?v=bFoUsDD_2ek (дата обращения 15.03.2012).

Контур руки, полученный из размытого бинаризованного изображения



Аппроксимация контура, уменьшение количества точек до сохранения общей формы контура



Вершины выпуклого многоугольника проверяются на остроту угла и расположение. Вершины острых углов, расположенные выше нижней границы блока, считаются найденными пальцами

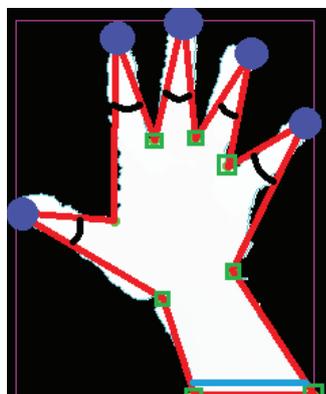


Рис. 6. Общий порядок распознавания пальцев на руке пользователя

7. Baak A., Müller M., Bharaj G., Seidel H., Theobalt C. A data-driven approach for real-time full body pose reconstruction from a depth camera // ICCV IEEE. – 2011, pp. 1092–1099.

References

1. *Xbox 360 Kinect Sensors and Bundles*, Available at: <http://www.microsoftstore.com> (accessed 14 January 2013).
2. *UDP: User Datagram Protocol*, Available at: <http://www.xserver.ru/computer/protokol> (accessed 14 January 2013).
3. *Drivers and libraries for the Xbox Kinect device on Windows*, Available at: <https://github.com/OpenKinect/libfreenect> (14 January 2013).
4. *Ocean Gesture Demo*, Available at: http://www.youtube.com/watch?v=gn_dnYbi4oE (accessed 11 March 2012).
5. *Jaguar Car Gesture Demo*, Available at: <http://www.youtube.com/watch?v=Utq6Z-JFSzk> (accessed 15 March 2012).
6. *Gesture Movie Control*, Available at: http://www.youtube.com/watch?v=bFoUsDD_2ek (accessed 25 March 2012).

7. Baak A., Müller M., Bharaj G., Seidel H., Theobalt C. A data-driven approach for real-time full body pose reconstruction from a depth camera // ICCV IEEE. 2011, pp. 1092–1099.

Рецензенты:

Гребенников В.И., д.ф.-м.н., профессор, главный научный сотрудник Института физики металлов УрО РАН, г. Екатеринбург;
 Готлиб Б.М., д.т.н., профессор, заведующий кафедрой «Мехатроника» ФГБОУ ВПО «Уральский государственный университет путей сообщения», г. Екатеринбург.
 Работа поступила в редакцию 15.01.2013.