

**Часть VI. «Переводные
статьи»**



Б. В. ЭЛСЕВИЕР (B. V. ELSEVIER),
ПЕРЕВОДЧИК: М. Р. ШАЛЕЛАШВИЛИ
РАСПОЗНАВАНИЕ ЖЕСТОВ РУК ДЛЯ
ВЗАИМОДЕЙСТВИЯ ЧЕЛОВЕКА
С КОМПЬЮТЕРОМ
[HAND GESTURE RECOGNITION FOR
HUMAN COMPUTER INTERACTION]

*7-АЯ МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ ПО РАЗВИТИЮ
В ВЫЧИСЛЕНИЯХ И КОММУНИКАЦИЯХ (2017)
[7TH INTERNATIONAL CONFERENCE ON ADVANCES
IN COMPUTING & COMMUNICATIONS, ICACC-2017, 22]
Индия, Кочин. 24 АВГУСТА 2017 ГОДА.
МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ,
2019*

Аннотация: *Использование физического контроллера, такого как мышь, клавиатура для взаимодействия человека с компьютером препятствует естественному интерфейсу, поскольку существует сильный барьер между пользователем и компьютером. В этой статье мы разработали надёжную систему распознавания жестов без маркеров, которая может эффективно отслеживать как статические, так и динамические жесты. Наша система переводит распознаваемый жест в такие действия, как открытие веб-сайтов и запуск приложений, таких как VLC Player и PowerPoint. Динамический жест используется для пролистывания слайдов в презентации. Наши результаты показывают, что интуитивное HCI (Human-computer interaction, человеко-машинное взаимодействие) может быть достигнуто с минимальными требованиями к оборудованию.*

Экспертиза проведена под руководством научного комитета 7-й Международной конференции по достижениям в области вычислительной техники и связи.

Peer-review under responsibility of the scientific committee of the 7th International Conference on Advances in Computing & Communications.

Ключевые слова: *человеко-машинный интерфейс, распознавание жестов, контуры, выпуклости, каскад Хаара.*

Keywords: *Human Computer Interaction; Hand gesture recognition; Contours; Convex hull; Haar cascade.*

1. ВСТУПЛЕНИЕ

Основная цель взаимодействия человека с компьютером заключается в улучшении взаимодействия между пользователями и компьютерами путём повышения восприимчивости компьютера к потребностям пользователей. Взаимодействие человека с персональным компьютером сегодня не ограничивается только взаимодействием клавиатуры и мыши. Взаимодействие между людьми происходит при помощи различных аспектов, таких как жест, речь и выражение лица. Способность взаимодействовать с системой естественным образом становится все более важной во многих областях взаимодействия человека с компьютером.

Для достижения распознавания жестов рук используются подходы, основанные как на не-видении, так и на видении. Примером подхода, не основанного на зрении, является обнаружение движения пальцев с помощью пары проводных перчаток. В целом подходы, основанные на зрении, более естественны, поскольку они не требуют ручных устройств. Теоретически в литературе жесты рук классифицируются на два типа: статические и динамические. Статические жесты рук можно определить как жесты, в которых положение и ориентация руки

в пространстве не изменяется в течение некоторого времени. Если есть какие-либо изменения в течение данного времени, жесты называются динамические жесты. Динамические жесты включают в себя такие жесты, как махание рукой, в то время как статические жесты включают соединение большого и указательного пальцев, чтобы сформировать символ “ОК”.

2. СУЩЕСТВУЮЩИЕ АНАЛОГИ

Проведённый обзор литературы даёт представление о различных методах, которые могут быть приняты и реализованы для достижения распознавания жестов рук. Это также помогает в понимании преимуществ и недостатков, связанных с различными техниками. Обзор литературы разделен на две основные фазы: модуль камеры и модуль обнаружения. Модуль камеры определяет различные камеры и маркеры, которые могут быть использованы. Модуль обнаружения общается с пре-процессингом изображений и извлечением характеристик.

Обычно используемые методы захвата ввода от наблюдаемого пользователя, — это перчатки данных, ручные ремни и камеры. Подход распознавания жестов [1] и [2] использует извлечение входных данных через перчатки данных. Для считывания движений рук используется ручной ремень с гироскопом, акселерометром и Bluetooth [3, 4]. Авторы [5] использовали креативную камеру Senz3D для захвата информации о цвете и глубине и [6] использовали стереокамеру Bumblebee2. Монокулярная камера была использована [7]. Экономичные модели, такие как [8, 9, 10], реализовали свои системы с использованием простых веб-камер. Методы [11, 12] используют камеру kinect depth RGB, которая использовалась для захвата цветового потока. Поскольку камеры глубины предоставляют дополнительную информацию о глубине для каждого пикселя (изображения глубины) с частотой кадров наряду с традиционными изображениями [13, 14]. В большинстве своём технологии позволяют надёжно извлечь информацию о положении руки, благодаря обработке представленной информации о цвете. Они не полностью решают основополагающую проблему. Эта фундаментальная проблема была решена в [15] при использовании черно-белого рисунка, дополненного реальными маркерами (монохромными точками опоры). Хотя встроенные веб-камеры не дают информации о глубине, они требуют меньших вычислительных затрат. Поэтому в нашей модели мы использовали веб-камеру, доступную в ноутбуке без использования каких-либо дополнительных камер или ручных маркеров, таких как перчатки.

Большое количество методов было использовано для предварительной обработки изображения, которое включает алгоритмы и методы для устранения шумов, обнаружения краёв, сглаживания с последующими различными методами

сегментации для извлечения границ, т. е. отделения переднего плана от заднего (фона). Авторы [9, 16] использовали алгоритм морфологии, который выполняет эрозию и расширение изображения для устранения шума. Для сглаживания контуров после бинаризации использовался гауссовский фильтр [10, 17]. Для выполнения сегментации в [6] была рассчитана карта глубины путём сопоставления левого и правого изображений с алгоритмом SAD (Sum of Absolute Differences, сумма абсолютных разностей). В [6] для поиска контуров использовался алгоритм Тео Павилдиса (Theo Pavlidis), который обрабатывает только граничащие друг с другом пиксели для нахождения контуров в изображении. Этот метод снижает вычислительные затраты. В [9, 13, 16] в качестве контура ладони руки был выбран самый большой контур, после которого контур был упрощён с использованием полигональной аппроксимации. Классификация — это процесс, в котором отдельные элементы группируются на основе сходства между элементами. Подход [18] использует классификатор на основе Евклидова расстояния для распознавания 25 поз рук. Классификатор машины поддержки вектора (SVM, Support Vector Machine) был использован в [19, 11]. Мы отклоняемся от других традиционных методов без использования каких-либо маркеров рук, таких как перчатки для распознавания жестов. В нашей модели мы использовали веб-камеру, доступную в ноутбуке без использования каких-либо дополнительных камер, сделав систему экономически эффективной. Таким образом, наша система находит применение в повседневной системе.

3. ПРЕДЛАГАЕМАЯ СИСТЕМА РАСПОЗНАВАНИЯ ЖЕСТОВ

Общая система состоит из двух частей: скрытой от пользователя (back-end) и видимой пользователю (front-end). Скрытая часть системы состоит из трёх модулей: модуль камеры, модуль обнаружения и модуль интерфейса, как показано на рис. 12. Их можно представить следующим образом:

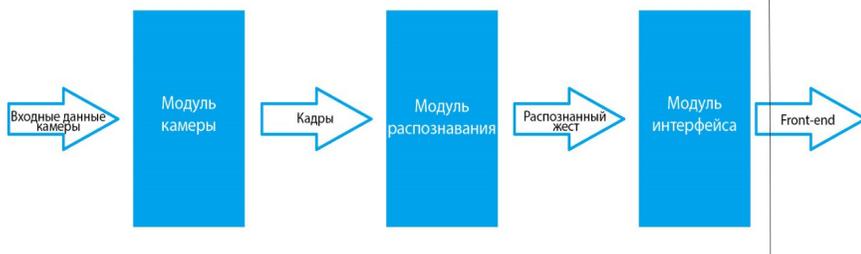


Рисунок 12 – Back-end архитектура

3.1 МОДУЛЬ КАМЕРЫ

Этот модуль отвечает за подключение и захват входной информации через различные типы детекторов изображения и отправляет это изображение в модуль обнаружения для обработки в виде кадров. Обычно используемые методы захвата ввода — это перчатки данных, ручные ремни и камеры. В нашей системе мы используем встроенную веб-камеру, которая экономична для распознавания как статических, так и динамических жестов. Система имеет соответствующее обеспечение для того, чтобы также позволить получить входную информацию и от USB веб-камеры, однако это потребует некоторых расходов от потребителя. Полученные кадры изображения выполнены в виде видео.

3.2 МОДУЛЬ ОБНАРУЖЕНИЯ

Этот модуль отвечает за обработку изображения. Выходная информация из модуля камеры подвергается различным методам обработки изображений, таким как преобразование цвета, удаление шума, сегментация изображения, после чего изображение подвергается экстракции контуров. Если изображение содержит дефекты, то на основе дефектов выпуклости затем определяются жесты. Если дефектов нет, то изображение классифицируется с помощью каскада Хаара для обнаружения жеста.

В случае динамических жестов, модуль обнаружения делает следующее; Если Microsoft PowerPoint был запущен с включённым слайд-шоу и веб-камера обнаружила ладонь в движении, для 5 последующих непрерывных кадров происходит определение жеста.

3.3 МОДУЛЬ ИНТЕРФЕЙСА (ВЗАИМОДЕЙСТВИЯ)

Этот модуль отвечает за сопоставление обнаруженных жестов рук с соответствующими действиями. Затем эти действия передаются соответствующему приложению. Видимая пользователю (front-end) часть состоит из трёх окон. Первое окно состоит из видеовхода, который снимается с камеры

с соответствующим названием обнаруженного жеста. Во втором окне отображаются контуры, найденные во входных изображениях. В третьем окне отображается сглаженная сегментированная версия изображения. Преимущество сложения окон с сегментами и контурами в качестве части графического пользовательского интерфейса заключается в том, чтобы пользователь знал о несоответствиях фона, которые повлияют на входную информацию в систему, и, таким образом, пользователь смог настроить свой ноутбук или настольную веб-камеру, чтобы избежать их. Это позволит добиться повышения производительности.

4. ПРОДВИГАЕМЫЙ МЕТОД

Мы предлагаем систему распознавания жестов без маркеров, которая следует очень эффективной методологии, как показано на рис. 13.

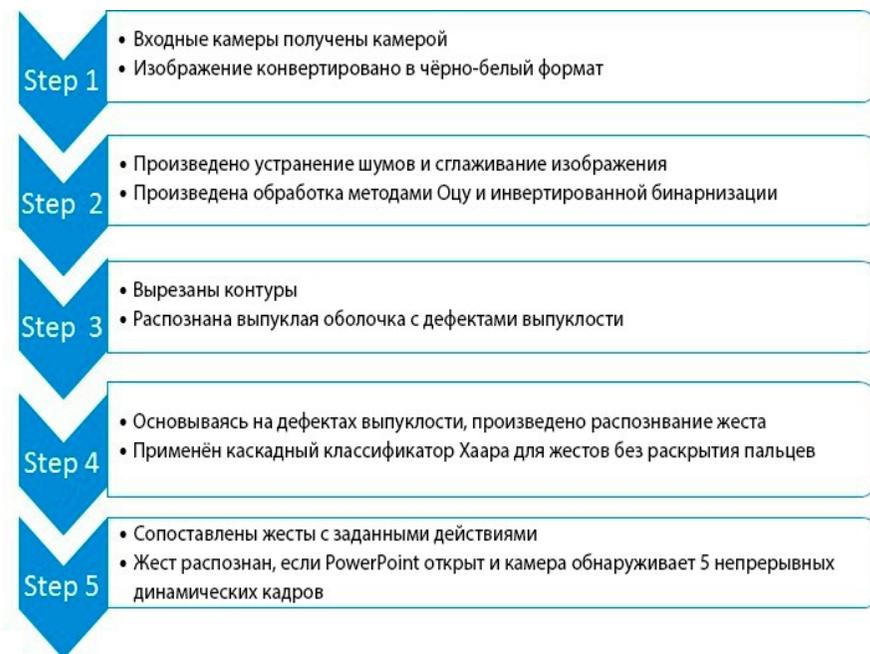


Рисунок 13 – Предложенный метод распознавания жестов

4.1. УСТРАНЕНИЕ ШУМОВ И СГЛАЖИВАНИЕ ИЗОБРАЖЕНИЯ

Входное изображение, которое находится в цветовом пространстве RGB, обрезается до размера 300 x 300 пикселей. Затем оно преобразуется в чёрно-белое изображение. Этот процесс показан на рис. 14.

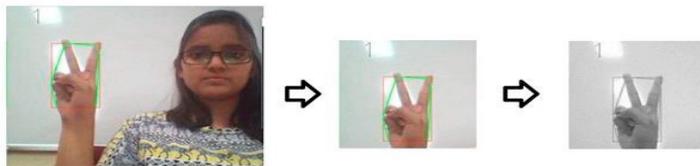


Рисунок 14 – Процесс обрезки и конвертации цветной картинку в чёрно-белую

Шум в изображениях можно определить как случайное отклонение яркости или цветовой информации, которая обычно создаётся в процессе получения изображения с веб-камеры. Этот шум является нежелательным аспектом изображения и должен быть удалён. С данной целью применяется размытие (фильтр Гаусса). Гауссовская фильтрация выполняется свёрткой гауссова ядра с каждой точкой входного массива. Затем они используются для создания выходного массива. 2D гауссово ядро может быть представлено математически, как показано в формуле 34.

$$G_0(x, y) = Ae^{-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(y-\mu_y)^2}{2\sigma_y^2}} \quad (34)$$

4.2. ПОРОГОВАЯ ОБРАБОТКА

Далее приступаем к пороговой обработке изображений (Image Thresholding), являющейся простым методом сегментирования изображения. Пороговая обработка применяется для получения двоичного изображения из чёрно-белого. Пороговый метод сравнивает каждое значение интенсивности пикселя (I) относительно порогового значения (T). Если $I < T$, конкретный пиксель заменяется чёрным пикселем, а если $I > T$, он заменяется белым пикселем. Грубо говоря, сводим все полутона к однозначному ответу «да» или «нет», есть тут пиксель или же его нет. В нашей работе используется пороговое значение (T) 127, которое классифицирует интенсивность пикселей в чёрно-белом изображении. Максимальное значение 255 — это значение пикселя, используемое, если какой-либо пиксель в изображении проходит пороговое значение. Два типа пороговой обработки (thresholding), которые мы применяем, являются инвертированным двоичным (бинарным) Thresholding и Otsu's Thresholding. Инвертированный двоичный thresholding инвертирует цвета,

чтобы получить белое изображением на чёрном фоне. Эта операция пороговой обработки может быть выражена, как показано в уравнении 35.

$$Dest(x, y) = \begin{cases} 0 & \text{if } src(x, y) > T \\ \max Val(255) & \text{otherwise} \end{cases} \quad (35)$$

Итак, если интенсивность пикселей $src(x, y)$ больше порогового значения T , то новая интенсивность пикселя инициализируется в 0. В противном случае пиксели устанавливаются в $\max Val$.

Нобуюки Оцу представил нам метод Оцу [20]. Благодаря этому методу достигается пороговое значение изображения на основе кластеризации. Бинаризация Otsu автоматически вычисляет пороговое значение из гистограммы изображения для бимодального изображения, которое является изображением, гистограмма которого имеет два пика. В методе Оцу мы пытаемся найти порог, который минимизирует внутриклассовую дисперсию (дисперсию внутри класса), определённую как взвешенную сумму дисперсий двух классов, как видно из формулы 36. Веса 0 и 1-вероятности двух классов, разделённых порогом t , а σ_0^2 и σ_1^2 — дисперсии. Вероятность класса 0, 1 (t) вычисляется из L гистограмм. Это показано в формуле 37.

$$\sigma_{\omega}^2(t) = \omega_0(t)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t) \quad (36)$$

$$\begin{aligned} \omega_0(t) &= \sum_{i=0}^{t-1} p(i) \\ \omega_1(t) &= \sum_{i=t}^{L-1} p(i) \end{aligned} \quad (37)$$

Метод Оцу показывает, что минимизация внутриклассовой дисперсии и максимизация межклассовой дисперсии генерируют те же результаты, что и показаны ниже в формуле 38.

$$\begin{aligned}\mu_0(t) &= \sum_{i=0}^{t-1} i \frac{p(i)}{\omega_0} \\ \mu_1(t) &= \sum_{i=t}^{L-1} i \frac{p(i)}{\omega_1} \\ \mu_T &= \sum_{i=0}^{L-1} ip(i)\end{aligned}\tag{39}$$

Следующие отношения в формуле 40 можно легко проверить.

$$\begin{aligned}\omega_0\mu_0 + \omega_1\mu_1 &= \mu_T \\ \omega_0 + \omega_1 &= 1\end{aligned}\tag{40}$$

Вероятности и способы класса могут вычисляться итеративно. Это может обеспечить эффективный алгоритм.

Прежде чем найти контуры, пороговая обработка была применена к двоичному изображению для достижения более высокой точности. Нижеприведенное изображение рис. 15 показывает окно переднего плана, отображающее пороговую версию ввода жестов пользователя.



Рисунок 15 – Front-end окно, отображающее жест после прохождения пороговой обработки

4.3. ВЫРЕЗАНИЕ КОНТУРОВ

Контуры являются полезным инструментом для обнаружения и распознавания объектов при обработке изображений. В нашей работе мы использовали контуры, чтобы обнаружить и распознать (отделить) руку от фона. Кривые, связывающие непрерывные точки одного цвета, называются контурами. Поиск контуров — это первый шаг, который похож на поиск белого объекта на чёрном фоне в OpenCV. Следовательно, инвертированное двоичное пороговое преобразование был использован во время порога. Второй шаг — нарисовать контуры, которые можно использовать для рисования любой формы при условии, что известны граничные точки. Некоторые жесты в нашей системе распознавания с соответствующими контурами показаны ниже на рис. 16.

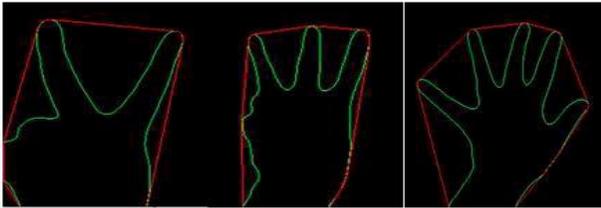


Рисунок 16 – Отделение контуров

4.4. ВЫПУКЛАЯ ОБОЛОЧКА И ДЕФЕКТЫ ВЫПУКЛОСТИ

Математически, выпуклая оболочка множества X точек в любом аффинном пространстве определяется как наименьшее выпуклое множество, содержащее X . Любое отклонение объекта от этой выпуклой оболочки можно рассматривать как дефект выпуклости. Выпуклая оболочка конечного точечного множества S может быть определена как множество всех выпуклых комбинаций его точек. В выпуклой комбинации каждой точке x_i в S присваивается вес α_i , и эти веса используются для вычисления среднего значения точек. Для каждого выбора веса полученная выпуклая комбинация является точкой в выпуклой оболочке. Выпуклая оболочка может быть представлена математически, как показано в формуле 41.

$$\text{Convex}(S) = \left\{ \sum_{i=1}^{|S|} \alpha_i x_i \mid (\forall i : \alpha_i \geq 0) \wedge \sum_{i=1}^{|S|} \alpha_i = 1 \right\} \quad (41)$$

4.5. КАСКАДНЫЙ КЛАССИФИКАТОР ХААРА

Для жестов, таких как ладонь и кулак, где нет дефектов выпуклости (пробелов между пальцами), используется классификатор каскада Хаара. Используется

коллекция положительных изображений, как минимум 10 оригинальных изображений, снятых при различных условиях освещения и углах. Каждое из исходных изображений обрезается, остаётся лишь интересующий объект. Коллекция негативных изображений, которая не содержит в себе интересующий объект, требует минимум 1000 изображений. Файл описания для отрицательных изображений создаётся с помощью библиотеки create samples. Каждое положительное изображение накладывается на минимум 200 изображений. Векторный файл создан на основе наложенных изображений (векторный файл должен содержать не менее 1500 снимков). Обучение Хаара будет использовать минимум 100 изображений размером 20 x 20, а обучение также может состоять из 15 или более этапов. Созданный XML-файл используется в качестве каскадного классификатора для обнаружения объектов в OpenCV.

5. ВНЕДРЕНИЕ И РЕЗУЛЬТАТЫ

В нашу систему распознавания жестов мы включили в общей сложности семь жестов, шесть из которых являются статическими, а один — динамическим. Эти статические жесты показаны ниже на рис. 17.

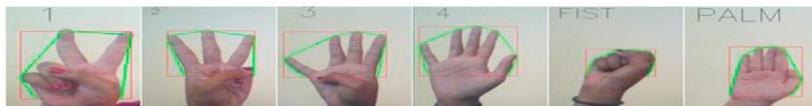


Рисунок 17 – Статические жесты в продвигаемой системе распознавания жестов.

Подписи, написанные в верхней части каждого жеста, т. е. «1», «2», обозначают количество дефектов выпуклости в каждом жесте. В жестах, которые не имеют никаких дефектов, т. е. кулак, ладонь, их название было написано как подпись над жестом.

Первый жест слева — знак «V» или знак «цифра 2», который запускает приложение VLC Media Player, как показано на рис. 40 (а). Второй — это жест «цифра 3», и он запускает домашнюю страницу Google в браузере пользователя по умолчанию, как показано на рис. 40 (б) и третий жест «цифра четыре», запускает домашнюю страницу YouTube. Четвёртый жест — это жест «цифра 5» или жест открытой ладони, который в нашей системе закрывает приложение, работающее на переднем плане. Пятый жест в приведённом выше изображении представляет собой сжатый кулак, который запускает Microsoft PowerPoint. Шестой и последний статический жест — это закрытая ладонь, которая переключает состояние (вкл./выкл.) Wi-Fi вычислительного устройства.

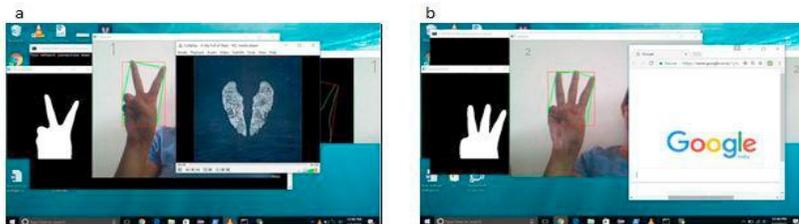


Рисунок 18 – (а) Жест “V” запускает VLC media player; (б) Жест “з” запускает стартовую страницу Google в браузере

В дополнение к вышеупомянутым статическим жестам модель также предусматривает динамический жест. Когда движущийся закрытый жест ладони распознаётся как 5 последовательных кадров, он считается динамическим движением «свайпа». Он используется, когда Microsoft PowerPoint работает на переднем плане, чтобы перейти к следующему слайду в презентации.

Наш первый подход к созданию системы распознавания жестов был методом вычитания фона. Вычитание фона, как следует из названия, представляет собой процесс отделения объектов переднего плана от фона в последовательности видеокадров. Это широко используемый подход для обнаружения движущихся объектов со статических камер. При реализации системы распознавания с использованием фонового вычитания мы столкнулись с рядом недостатков и проблем неточности. Вычитание фона не срабатывает корректно в случаях с внезапными, резкими изменениями освещения, ведущими к нескольким несоответствиям. Этот метод также требует сравнительно много параметров, которые нужно подбирать с умом. Из-за этих осложнений мы приняли решение использовать контуры, дефекты выпуклости и каскад Хаара для обнаружения объекта (руки). Сочетание этих методов позволило нам достичь большей точности и преодолеть проблемы, с которыми мы столкнулись при использовании отделения объекта от фона.

Чтобы оценить точность нашей системы, мы провели два набора оценок. В первом наборе оценок мы использовали пространства, которые содержали различные виды простых фонов без каких-либо несоответствий. Во второй оценке мы использовали задние фоны с несколькими несоответствиями. Каждый жест был выполнен 10 раз в каждом из заданных условий. Среднее число раз, когда тот или иной жест был распознан правильно, было принято за его точность в процентах, а полученная точность показана в таблице 7. При реализации на любом простом фоне система распознавания жестов была надёжной и распознавание проводилось с хорошей точностью. Эта точность сохранялась независимо от цвета фона при условии, что он представляет собой простой сплошной фон, лишенный каких-либо несоответствий.

В тех случаях, когда фон не был простым, объекты на заднем фоне привели к ошибкам в процессе захвата изображения, что привело к неисправным выходам.

Таким образом, точность была не так хороша, как в сценариях с простым фоном. После наблюдения за результатами, полученными системой распознавания жестов с различными условиями заднего плана, рекомендуется использовать эту систему с простым однородным фоном для получения наилучших результатов и высокой точности.

Таблица 7 – Точность распознавания каждого жеста с однотонным и неоднотонным фоном

Жест	Точность с однотонным фоном (в %)	Точность с неоднотонным фоном (в %)
«2 пальца» (1 дефект выпуклости)	94	40
«3 пальца» (2 дефекта выпуклости)	93	50
«4 пальца» (3 дефекта выпуклости)	92	48
«5 пальцев» (4 дефекта выпуклости)	92	52
Ладонь	95	92
Кулак	95	92
Прокрутка/свайп (динамический жест)	85	80

6. ЗАКЛЮЧЕНИЕ И БУДУЩАЯ РАБОТА

Мы смогли создать надёжную систему распознавания жестов, не использующую никаких маркеров, что сделало её более удобной и недорогой. В этой системе распознавания жестов мы стремились обеспечить жесты, охватывающие почти все аспекты человеко-машинного взаимодействия (HCI), такие, как функциональные возможности системы, запуск приложений и открытие некоторых популярных веб-сайтов. В будущем мы хотели бы улучшить точность и добавить больше жестов для реализации большего количества функций. Наконец, мы намерены расширить наши базовые сценарии и применить наш механизм отслеживания в различных аппаратных средствах, включая цифровое телевидение и мобильные устройства. Мы также стремимся распространить этот механизм на ряд пользователей, включая инвалидов.

ЛИТЕРАТУРА

1. Granit Luzhnica, Elizabeth Lex, Viktoria Pammer. A Sliding Window Approach to Natural Hand Gesture Recognition using a Custom Data Glove. In: 3D User Interfaces (3DUI); 2016 IEEE Symposium on; 2016 Mar 19; New York: IEEE; 2016; p.81-90.
2. Ji-Hwan Kim, Nguyen Duc Thang, Tae-Seong Kim. 3-D hand Motion Tracking and Gesture Recognition Using a Data Glove. In: Industrial Electronics; 2009 IEEE International Symposium on; 2009 July 5; New York: IEEE;2009; p.1013-1018.
3. Hung CH, Bai YW, Wu HY. Home outlet and LED array lamp controlled by a smartphone with a hand gesture recognition. In: Consumer Electronics (ICCE); 2016 IEEE International Conference on; 2016 Jan 7; New York: IEEE;2016; p.5-6.
4. Hung CH, Bai YW, Wu HY. Home appliance control by a hand gesture recognition belt in LED array lamp case. In: Consumer Electronics (GCCE); 2015 IEEE 4th Global Conference on; 2015 Oct 27; New York: IEEE;2015; p. 599-600
5. She Y, Wang Q, Jia Y, Gu T, He Q, Yang B. A real-time hand gesture recognition approach based on motion features of feature points. In: Computational Science and Engineering (CSE); 2014 IEEE 17th International Conference on; 2014 Dec 19; New York: IEEE;2014; p.10961102.
6. Lee DH, Hong KS. A Hand gesture recognition system based on difference image entropy. In: Advanced Information Management and Service (IMS), 2010 6th International Conference on; 2010 Nov 30; Seoul; New York: IEEE; 2010; p. 410-413.
7. Dulayatrakul J, Prasertsakul P, Kondo T, Nilkhamhang I. Robust implementation of hand gesture recognition for remote human-machine interaction. In: Information Technology and Electrical Engineering (ICITEE); 2015 7th International Conference on;2015 Oct 29; p. 247252.
8. Tsai TH, Huang CC, Zhang KL. Embedded virtual mouse system by using hand gesture recognition. In: Consumer Electronics-Taiwan (ICCE-TW); 2015 IEEE International Conference on; 2015 Jun 6; Taiwan, Taipei; New York: IEEE; 2015; p. 352-353.
9. Hussain I, Talukdar AK, Sarma KK. Hand gesture recognition system with real-time palm tracking. In: India Conference (INDICON);2014 Annual IEEE;2014 Dec 11; India, Pune; New York: IEEE; 2014; p. 1-6.
10. Huong TN, Huu TV, Le Xuan T. Static hand gesture recognition for vietnamese sign language (VSL) using principle components analysis. In: Communications, Management and Telecommunications (ComManTel); 2015 International Conference on; 2015 Dec 28; p. 138-141.
11. Chen Y, Luo B, Chen YL, Liang G, Wu X. A real-time dynamic hand gesture recognition system using kinect sensor. In: Robotics and Biomimetics (ROBIO); 2015 IEEE International Conference on; 2015 Dec 6; New York: IEEE;2015; p. 2026-2030.
12. C. Wang, Z. Liu and S. C. Chan. Superpixel-Based Hand Gesture Recognition With Kinect Depth Camera. *IEEE Transactions on Multimedia* 2015; **17**(1): 29-39.
13. Chen WL, Wu CH, Lin CH. Depth-based hand gesture recognition using hand movements and defects. In: Next-Generation Electronics (ISNE); 2015 International Symposium on; 2015 May 4; Taiwan, Taipei; New York: IEEE;2015; p. 1-4.
14. Wong WS, Hsu SC, Huang CL. Virtual touchpad: Hand gesture recognition for smartphone with depth camera. In: Consumer ElectronicsTaiwan (ICCE-TW);2015 IEEE International Conference on; 2015 Jun 6; Taiwan, Taipei; New York: IEEE;2015; p. 214-215.
15. Ishiyama H, Kurabayashi S. Monochrome glove: A robust real-time hand gesture recognition method by using a fabric glove with design of structured markers. In: Virtual Reality (VR); 2016 IEEE;2016 Mar 19; Greenville, SC; New York: IEEE;2016; p. 187-188.

16. Suriya R, Vijayachamundeeswari V. A survey on hand gesture recognition for simple mouse control. In: Information Communication and Embedded Systems (ICICES); 2014 International Conference on; 2014 Feb 27; India, Chennai; New York: IEEE;2014; p. 1-5.
17. Chanda K, Ahmed W, Mitra S. A new hand gesture recognition scheme for similarity measurement in a vision based barehanded approach. In: Image Information Processing (ICIIP); 2015 Third International Conference on; 2015 Dec 21;; New York: IEEE;2015; pp. 17-22.
18. Luzhnica G, Simon J, Lex E, Pammer V. A sliding window approach to natural hand gesture recognition using a custom data glove. In:3D User Interfaces (3DUI); 2016 IEEE Symposium on; 2016 Mar 19; Greenville, SC; New York: IEEE;2016; p. 81-90.
19. Chen Y, Ding Z, Chen YL, Wu X. Rapid recognition of dynamic hand gestures using leap motion. In: Information and Automation; 2015 IEEE International Conference on; 2015 Aug 8; New York: IEEE;2015; p. 1419-1424.
20. Otsu, Nobuyuki. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics* 1979; **9**(1): 62-66.

