

УДК 004.5

**МЕТОД СЕГМЕНТАЦИИ РУК ДЛЯ РАСПОЗНАВАНИЯ ЖЕСТОВ РУК
ДЛЯ ЕСТЕСТВЕННОГО ВЗАИМОДЕЙСТВИЯ ЧЕЛОВЕКА И
КОМПЬЮТЕРА****Савушкина Анастасия Владимировна**студент КФ МГТУ им. Н.Э. Баумана,
248000, РФ, г. Калуга, ул. Баженова, дом № 2
e-mail: nastia-sava1999@yandex.ru**Федоров Виктор Олегович**кандидат технических наук, доцент КФ МГТУ им. Н.Э. Баумана,
248000, РФ, г. Калуга, ул. Баженова, дом № 2
e-mail: fedorov_vo@bmstu.ru**Аннотация**

При взаимодействии человека и компьютера одним из основных средств бесконтактной связи являются жесты рук. Однако существуют проблемы при внедрении систем распознавания жестов на основе компьютерного зрения из-за высоких требований к точности и скорости реакции. Целью данной статьи является рассмотрение одного из самых быстрых методов распознавания жестов рук в реальном времени, основанного на сегментации кисти руки. Распознавание жестов рук имеет большое значение во многих приложениях, таких как распознавание языка жестов, дополненная реальность (виртуальная реальность), сурдоперевод и управление роботами.

Ключевые слова: жест, распознавание, управление, сегментация, классификация изображений, сурдоперевод, человеко-машинное взаимодействие.

**HAND SEGMENTATION METHOD FOR RECOGNIZING HAND GESTURES
FOR NATURAL HUMAN-COMPUTER INTERACTION****Anastasia V. Savushkina**Student of Bauman Moscow State Technical University (Kaluga Branch), 248000, Russian Federation, Kaluga, Bazhenova st., 2.
nastia-sava1999@yandex.ru**Viktor O. Fedorov**Candidate of Technical Sciences, Docent of Bauman Moscow State Technical University (Kaluga Branch), 248000, Russian Federation, Kaluga, Bazhenova st., 2.
fedorov_vo@bmstu.ru

ABSTRACT

When a person and a computer interact, one of the main means of contactless communication is hand gestures. However, there are problems with the introduction of gesture recognition systems based on computer vision due to the high requirements for accuracy and speed of reaction. The purpose of this article is to consider one of the fastest real-time hand gesture recognition methods based on hand segmentation. Hand gesture recognition is of great importance in many applications, such as sign language recognition, augmented reality (virtual reality), sign language translation and robot control.

Keywords: gesture, recognition, control, segmentation, image classification, sign language translation, human-machine interaction.

Введение

Взаимодействие человека с компьютером является неотъемлемым аспектом современного мира. Однако, к сожалению, привычные технологии взаимодействия с техникой имеют ряд недостатков, которые не позволяют комфортно пользоваться ими людям с ограниченными возможностями. Так, например, люди с инвалидностью по слуху и речи не могут посылать голосовые команды, которые будут интерпретированы машиной, поэтому появилась необходимость в развитии новой технологии управления периферийными устройствами. Для успешного, качественного и быстрого взаимодействия между субъектами такая технология, должна будет функционировать в режиме реального времени [2].

Если невозможно использовать другие средства коммуникации, остается использовать невербальные способы общения, например жестовое. Жест является наиболее простым и естественным способом общения, который помогает передавать информацию и более эффективно выражать смысл мыслей и чувств. Для знакомых с сурдоалфавитом людей этот способ общения является интуитивным и естественным средством связи между людьми. Вместе с тем, они также могут послужить средством взаимодействия между людьми и машинами.

В данной статье будет описан метод сегментации кисти руки для распознавания жестов. Существующие технологии управления в основном требуют использования специальных устройств для обнаружения основных точек руки. Предлагаемый метод не предполагает использование посредников для реализации управления, а также распознавание жеста основывается на выделении траектории движения отдельных ее фрагментов, что упрощает и ускоряет один из важных этапов - процесс классификации.

Этапы метода распознавания жестов рук.

Распознавание жестов рук можно разделить на четыре этапа. Первый этап отвечает за выявление области кисти руки методом вычитания фона: результат представляется в преобразованной в двоичное изображение форме. На втором этапе выполняется сегментация ладони и пальцев, чтобы упростить их распознавание. Обнаружение и распознавание пальцев происходит на третьем этапе. И на последнем четвертом этапе с помощью простого классификатора правил распознаются жесты рук. Данные этапы схематично представлены на рисунке 1.

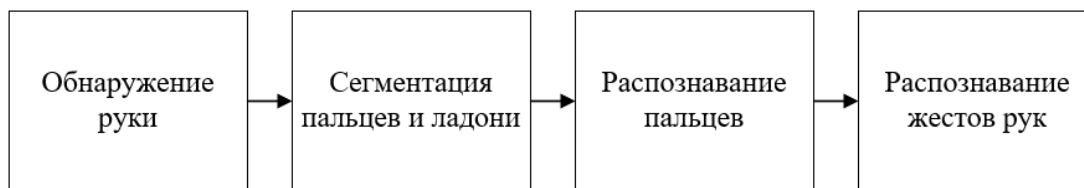


Рисунок 1 – Этапы метода распознавания жестов рук

В качестве устройств ввода применяются три типа камер – обычная, стерео и ToF (time of flight). Первый тип камер широко распространен, при его использовании цвет кожи и характерные точки объединяются для более точной сегментации области кисти [6]. Изображения из стерео и ToF камеры содержат информацию о глубине, что позволяет, используя карту глубины изображения, легко отделить область руки от фона.

Метод вычитания фона является простым и действенным способом для того, чтобы определить область кисти руки на исходном изображении [5]. Тем не менее в отдельных случаях данный метод захватывает и другие движущиеся объекты помимо искомого. Чтобы минимизировать случаи некорректного определения области кисти руки, используется палитра оттенков кожи человека, заранее добавленная в базу знаний о пользователе. Изменение размера изображения обнаруженной руки зависит от инвариантности распознавания жестов к масштабу изображения.

Сегментация пальцев и ладони выполняется на двоичном изображении, полученном в результате обнаружения кисти руки и содержащем два типа пикселей. Белые пиксели отвечают за элементы области кисти руки и черные пиксели, которые принадлежат фону. Необходимые параметры для сегментации пальцев, изображенные на рисунке 2, были взяты из статьи Хуэйси Ву и др. [9] и модифицированы в соответствии с текущей задачей.

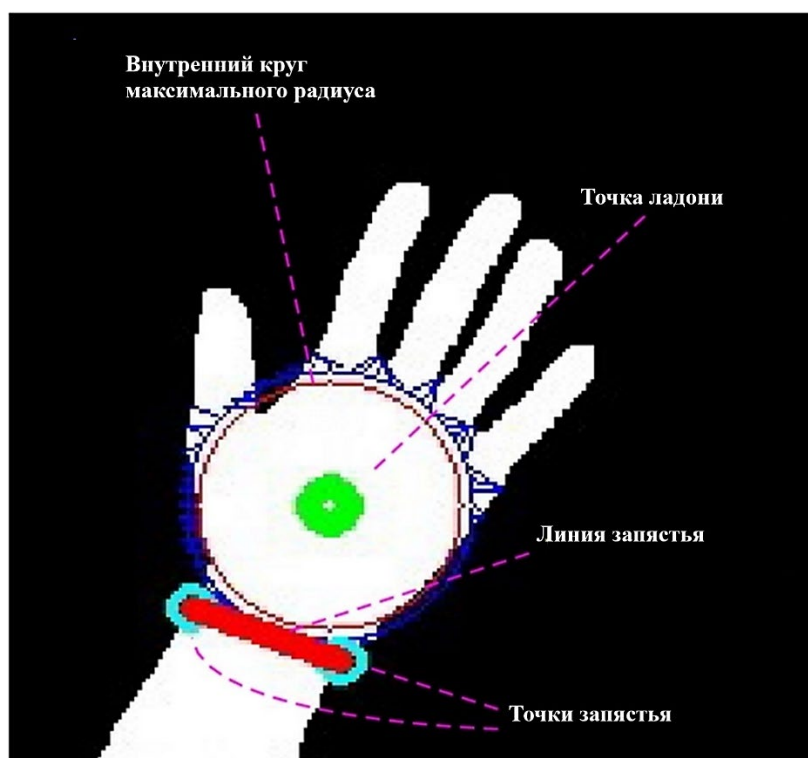


Рисунок 2 - Точка ладони, точки запястья, линия запястья и внутренний круг максимального радиуса.

Сегментацию следует начинать с определения центральной точки ладони с помощью метода карты расстояний [4]. Суть данного метода заключается в том, что каждому пикселю

изображения добавляется расстояние от него до ближайшего граничного пикселя. В результате преобразования расстоянием центральной точкой ладони является пиксель с наибольшим расстоянием.

Внутренним кругом называется окружность, проходящая через крайнюю точку с центром в центральной точке ладони и находящаяся внутри ладони. Увеличение радиуса круга происходит до тех пор, пока не будет достигнут край ладони, который содержит черные пиксели.

Маска ладони формируется из связанных ближайших граничных точек. Для этого строится окружность радиусом в 1,2 раза больше радиуса максимального внутреннего круга [9], на которой выбираются равномерно распределённые точки. Важно понимать, что пиксели относятся только к двоичному изображению, в то время как точки – это метки, которые необходимы для сегментации, или вспомогательные элементы. Для каждой выбранной точки на окружности определяется ближайшая к ней граничная точка и выравнивается по ней. Точка называется граничной, если восемь соседей пикселя состоят из белых и черных пикселей. Увеличенный радиус максимального внутреннего круга используется для формирования более точной маски ладони. Маску ладони можно использовать для легкой сегментации пальцев, потому что она полностью закрывает часть руки, не отвечающую за пальцы. Результат сегментации пальцев показан на рисунке 3.



Рисунок 3. – Результат сегментации пальцев

Важными точками для распознавания жестов рук являются две точки на концах линии запястья, проходящей через нижнюю часть кисти. Точками запястья называются две последовательные точки маски, расстояние между которыми максимально.

Чтобы отследить вращение кисти руки, на двоичном изображении строится стрелка, исходящая из точки ладони к средней точке линии запястья, и настраивается в направлении севера. Жест руки должен быть инвариантным к вращению, поэтому изображение руки вращается синхронно.

Определение вида распознанного пальца

Области пальцев на сегментированном изображении определяются с помощью алгоритма маркировки [8]. Пальцами считаются области, размер которых достаточен для покрытия пальца, если количество пикселей мало, то область считается зашумленной и отбрасывается. Вокруг каждой области пальца формируется минимальная

ограничивающая рамка, центр которой используется в качестве центральной точки пальца. Большой палец обнаруживается и распознается отдельно от других пальцев.

Наличие большого пальца определяется по градусу угла между линией, соединяющей центр пальца с точкой ладони, и линией запястья. Если существует градус меньше 50° , то на изображении руки есть большой палец. Центральной точкой большого пальца будет соответствующий центр минимальной ограниченной рамки. Если все градусы больше 50° , то большого пальца на изображении нет.

Для распознавания и обнаружения других пальцев необходимо определить линию ладони, параллельную линии запястья. Алгоритм поиска линии ладони начинается с ряда линии запястья. Линия, пересекающая кисть и параллельная линии запястья, есть в каждом ряду. В случае, если большой палец обнаружен, линия смещается вверх с краевыми точками ладони вместо большого пальца в качестве начальной точки линии. Смещение линии зависит от количества связанных наборов белых пикселей на пересечении линии и руки. Если имеется более одного набора, то линия рассматривается как кандидат на линию ладони, а если такой набор один, то происходит смещение линии [7]. Однако, если на изображении руки не обнаружен большой палец, линией ладони является линия, пересекающая руку с более чем одним связанным набором белых пикселей в их пересечении. Когда большой палец отсутствует, на пересечении линии и руки остается только один связанный набор пикселей, поэтому линия ладони будет найдена, как только количество связанных наборов белых пикселей станет равным двум.

Полученная линия ладони делится на четыре части, к которым определяются, в соответствии с горизонтальной координатой центральной точки, пальцы. Палец попавший в первую часть является указательным пальцем. Палец, который относится ко второй части, будет средним пальцем. Безымянный палец и мизинец соответственно относятся к третьей и четвертой части. Результат распознавания пальцев показан на рисунке 4.

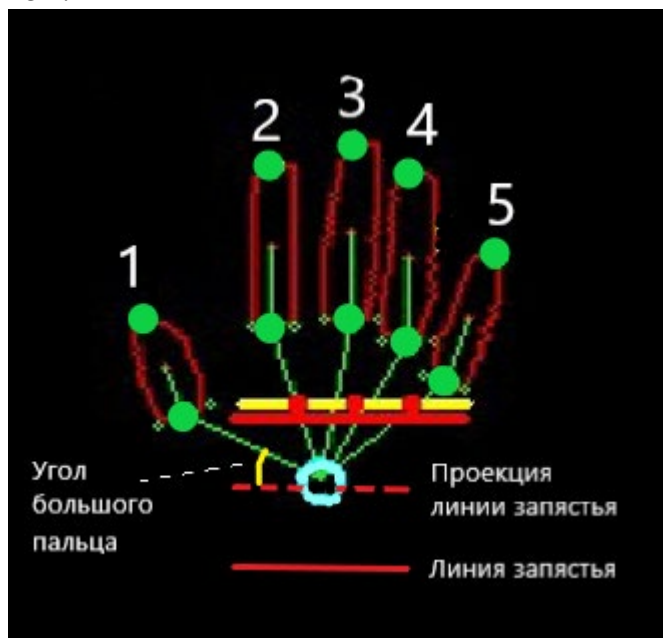


Рисунок 4. – Линия ладони

В случаях, когда между пальцами отсутствует промежуток из-за близкого расположения двух и более пальцев, используется ширина минимальной ограничивающей рамки в качестве показателя разграничения. Обнаруженная область описывает один палец, если ширина минимальной ограничивающей рамки равна обычному значению. В случае превышения обычного значения в несколько раз, обнаруженная область содержит несколько пальцев, которые расположены плотно к друг другу. Чтобы распознавание

пальцев было точным, используются данные о расстояниях и углах между пальцами для различных жестов.

Простой классификатор правил применяется для распознавания жестов, после того как все пальцы были обнаружены и распознаны. Классификатор правил очень эффективен и действенен. Жест руки предсказывается исходя из количества и наименований обнаруженных пальцев [3]

Заключение

В этой статье был рассмотрен метод распознавания жестов рук с использованием сегментации пальцев, который хорошо работает в режиме реального времени. Результат обнаружения руки играет значимую роль в эффективности выполнения предлагаемого метода. Алгоритмы машинного обучения способны отличить руку от фона [1]. Тем не менее качество распознавания может ухудшаться при наличии движущихся объектов с цветом, который похож на цвет кожи, потому что метод вычитания фона может включить их в область обнаруженной руки. Совместное использование стерео- и ToF камер, которые передают информацию о глубине, с методами машинного обучения в будущем способно решить проблему обнаружения руки, связанную с фоном, а также повысить качество и надежность.

Список литературы:

1. Ростовцев, В. С. Искусственные нейронные сети : учебник для вузов / В. С. Ростовцев. – 2-е изд., стер. – Санкт-Петербург: Лань, 2021. – 216 с. – ISBN 978-5-8114-7462-2. – Текст: электронный // Лань: электронно-библиотечная система. – URL: <https://e.lanbook.com/book/160142> (дата обращения: 15.10.2022).
2. Ян, Э. С. Программирование компьютерного зрения на языке Python / Э. С. Ян ; перевод с английского А. А. Слинкин. – Москва: ДМК Пресс, 2016. – 312 с. – ISBN 978-5-97060-200-3. – Текст: электронный // Лань: электронно-библиотечная система. – URL: <https://e.lanbook.com/book/93569> (дата обращения: 15.10.2022).
3. Селянкин, В. В. Компьютерное зрение. Анализ и обработка изображений : учебник для вузов / В. В. Селянкин. – 2-е изд., стер. – Санкт-Петербург: Лань, 2021. – 152 с. – ISBN 978-5-8114-8259-7. – Текст: электронный // Лань: электронно-библиотечная система. – URL: <https://e.lanbook.com/book/173806> (дата обращения: 15.10.2022).
4. Segmentation of 3d meshes combining the artificial neural network classifier and the spectral clustering / F.R. Zakani, M. Bouksim, K. Arhid [и др.] // Компьютерная оптика. – 2018. – № 2. – С. 312-319. – ISSN 0134-2452. – Текст: электронный // Лань: электронно-библиотечная система. – URL: <https://e.lanbook.com/journal/issue/308468> (дата обращения: 15.10.2022).
5. Meenu Sahu, Amit Saxena, Manish Manoria, Application of Feature Extraction Technique: A Review. // International Journal of Computer Science and Information Technologies, pp. 3014-3016, Vol 6 (3), 2015. – URL: <http://ijcsit.com/docs/Volume%206/vol6issue03/ijcsit20150603227.pdf> (дата обращения: 15.10.2022)
6. Шапиро, Л. Компьютерное зрение: учебное пособие / Л. Шапиро, Д. Стокман ; под редакцией С. М. Соколова ; перевод с английского А. А. Богуславского. – 4-е изд. – Москва: Лаборатория знаний, 2020. – ISBN 978-5-00101-696-0. – Текст:

- электронный // Лань: электронно-библиотечная система. – URL: <https://e.lanbook.com/book/135496> (дата обращения: 15.10.2022).
7. Пролубников, А. В. Математические методы распознавания образов: учебное пособие / А. В. Пролубников. – Омск: ОмГУ, 2020. – 110 с. – ISBN 978-5-7779-2461-2. – Текст: электронный // Лань: электронно-библиотечная система. – URL: <https://e.lanbook.com/book/142454> (дата обращения: 15.10.2022).
8. Цуриков, А. Н. Моделирование и обучение искусственных нейронных сетей: учебное пособие / А. Н. Цуриков. – Ростов-на-Дону: РГУПС, 2019. – 112 с. – ISBN 978-5-88814-867-9. – Текст: электронный // Лань: электронно-библиотечная система. – URL: <https://e.lanbook.com/book/140610> (дата обращения: 15.10.2022).
9. Huisi Wu, Lei Wang, Mingjun Song, Zhengkun Wen, Automatic Hand Gesture Recognition Based on Shape Context. // Foundations of Intelligent Systems, pp. 889-900, 2014. - URL: https://www.researchgate.net/publication/278687025_Automatic_Hand_Gesture_Recognition_Based_on_Shape_Context (дата обращения 15.10.2022)

References:

1. Rostovtsev, V. S. Artificial neural networks : textbook for universities / V. S. Rostovtsev. – 2nd ed., erased. – Saint Petersburg: Lan, 2021. – 216 p. – ISBN at 978-5-8114-7462-2. – Text: electronic // Lan: electronic library system. – Address: <https://e.lanbook.com/book/160142> (accessed: 10/15/2022).
2. Yan, E. S. Programming computer vision in Python in / E. S. Yan ; translated from English by A. A. Slinkin. – Moscow: DMK Press, 2016. – 312 p. – ISBN at 978-5-97060-200-3. – Text: electronic // Lan: electronic library system. – Address: <https://e.lanbook.com/book/93569> (date of application: 10/15/2022).
3. Selyankin, V. V. Computer vision. Image analysis and processing : textbook for universities / V. V. Selyankin. – 2nd ed., erased. – Saint Petersburg: Lan, 2021. – 152 p. – ISBN at 978-5-8114-8259-7. – Text: electronic // Lan: electronic library system. – Address: <https://e.lanbook.com/book/173806> (accessed: 10/15/2022).
4. Segmentation of a 3D grid combining artificial neural network classifier and spectral clusters / F. R. Zakani, M. Buksim, K. Arhid [et al.] // Computer optics. – 2018. – No. 2. – pp. 312-319. – ISSN 0134-2452. – Text: electronic // Lan: electronic library system. – Address: <https://e.lanbook.com/journal/issue/308468> (accessed: 10/15/2022).
5. Minu Sahu, Amit Saxena, Manish Manoria, Application of the method of feature extraction: An overview. // International Journal of Computer Science and Information Technology, pp. 3014-3016, Volume 6 (3), 2015. – URL: <http://ijcsit.com/docs/Volume%206/vol6issue03/ijcsit20150603227.pdf> (date of request: 10/15/2022)
6. Shapiro, L. Computer Vision: textbook / L. Shapiro, D. Stockman ; edited by S. M. Sokolov; translated from English by A. A. Boguslavsky. – 4th ed. – Moscow: Laboratory of Knowledge, 2020. – ISBN at 978-5-00101-696-0. – Text: electronic // Lan: electronic Library the system. – Address: <https://e.lanbook.com/book/135496> (accessed: 10/15/2022).

7. Prolubnikov, A.V. Mathematical methods of pattern recognition: a textbook / A.V. Prolubnikov. – Omsk: OmSU, 2020. – 110 p. – ISBN at 978-5-7779-2461-2. – Text: electronic // Lan: electronic library system. – Address: <https://e.lanbook.com/book/142454> (accessed: 10/15/2022).
8. Tsurikov, A. N. Modeling and training of artificial neural networks: a textbook / A. N. Tsurikov. – Rostov-on-Don: RSUPS, 2019. – 112 p. – ISBN at 978-5-88814-867-9. – Text: electronic // Lan: electronic library system. – Address: <https://e.lanbook.com/book/140610> (accessed: 10/15/2022).
9. Huixi Wu, Lei Wang, Mingjun Song, Zhengkun Wen, Automatic recognition of hand gestures based on the context of the form. // Fundamentals of Intelligent Systems, pp. 889-900, 2014. - URL: https://www.researchgate.net/publication/278687025_Automatic_Hand_Gesture_Recognition_Based_on_Shape_Context (accessed 15.10.2022)