

## ИНФОРМАТИКА

doi: 10.51639/2713-0576\_2023\_3\_2\_180

УДК 004.021+004.93

ГРНТИ 20.53.19

ВАК 05.13.17

### Алгоритм распознавания жестов руки на основе метода динамической цветовой сегментации видеокadra

\* Сингх С., Прибыльский А. В.

*ЮФУ, Институт радиотехнических систем и управления  
347922, Ростовская обл., Таганрог, Некрасовский пер. 42*

email: \* [singkh@sfedu.ru](mailto:singkh@sfedu.ru), [apribylsky@sfedu.ru](mailto:apribylsky@sfedu.ru)

В представленной работе рассматривается задача сегментации объектов на видеоизображении с использованием динамической цветовой сегментации видеокadra для дальнейшей классификации сегментированных объектов. Метод динамической цветовой сегментации использует алгоритмы калибровки цвета, извлечения информации о цветовых свойства объекта на видео, предобработку видеокadra, затем создает морфологически обработанную маску и применяет её к входящему видеопотоку. Результирующая маска формируется на основании двух масок сформированных из изображений в цветовых пространствах HSV и YCBCR. Целью является исследование методов цветовой сегментации видеокadra и проведение экспериментов, для использования в системе распознавания жестов руки. В ходе экспериментов составлен алгоритм динамической сегментации изображения, получены изображения с результатами сегментации с использованием статических и динамических методов цветовой сегментации изображения, приведен график сравнения работы статически и динамической сегментации при различных уровнях яркости, приведено описание полученных изображений и графиков, сформулирован вывод по результатам эксперимента.

*Ключевые слова:* цветровая сегментация, MatLab, видеокادر, фильтрация, морфологическая обработка.

### Введение

Современные технологии компьютерного зрения и распознавания жестов руки позволяют создавать инновационные системы управления, которые обладают высокой степенью точности и удобства в использовании. Однако, для достижения таких результатов необходимо разрабатывать системы и алгоритмы работы учитывающие изменения, как во внешней среде, так и во внутренней. Системы распознавания объектов в большинстве случаев работают с видеоизображения, которое представляет собой непрерывно следующий друг за другом набор видеокadров. Т. к. видеокادر это по своей сути изображения, то к нему применяют соответствующую методологию. Одним из основных этапов в распознавании жестов руки является сегментация, видеокadra, то есть выделения изображения руки на фоне, которое может осуществляться различными методами, например методом динамической цветовой сегментации, которая основывается на изменении цветового порога в зависимости от освещения и цвета кожи руки. Этот метод позволяет выделить руку на

видеоизображения с учётом изменяющегося цвета и контекста. После чего с помощью нейронного детектора и классификатора можно определить, к какому классу относится распознанный жест и есть ли он вообще в кадре.

### Теория и методы исследования

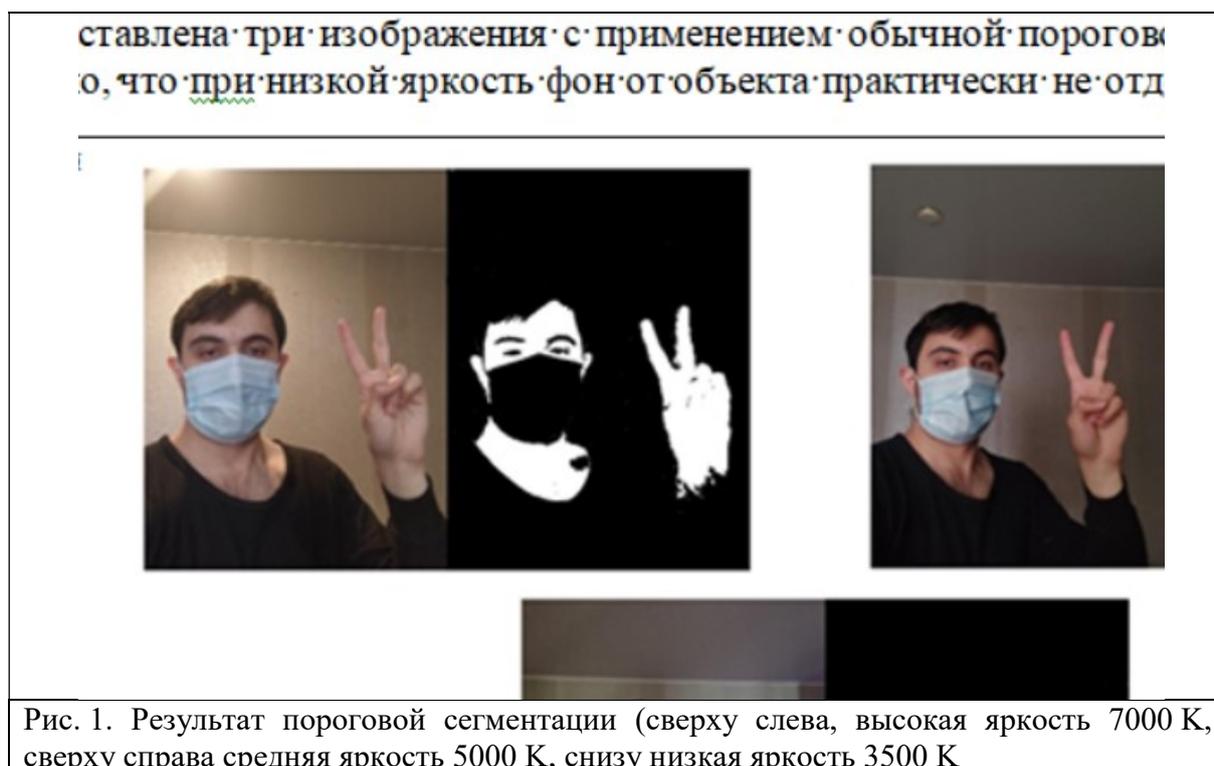
Цветовая сегментация является процессом разделения изображения на отдельные части, которые соответствуют объектам или областям с различными цветами. Вот несколько видов цветовой сегментации:

1. По пороговому значению: пороговая сегментация является самой простой формой цветовой сегментации, где пиксели сравниваются с определенным пороговым значением, чтобы разделить их на две категории – фон и объект. Эта техника проста в реализации и обработке, но не очень точна для сложных изображений.

2. К-средних: это метод кластеризации, используемый для сегментации изображений на основе цветов. Он работает путём разделения всех пикселей на заданное количество групп (K) с использованием алгоритма минимизации расстояния между пикселями в каждой группе. Этот метод даёт более точные результаты, чем пороговая сегментация, но может быть несколько более сложным для реализации.

3. Алгоритм разреза на графах: этот метод использует технику графического разреза для разделения изображений на отдельные сегменты. Он основан на поиске минимального среза в графе, который разделяет пиксели на заданные сегменты. Этот метод может быть достаточно точным, но может потребовать больше времени для обработки изображений высокого разрешения.

Пороговая сегментация не требует высоких вычислительных затрат, в отличие от K-средних и алгоритма разреза, а невысокую точность можно компенсировать динамическим порогом и обработкой изображения. Обычная пороговая сегментация, чувствительна не только к цветовому порогу, но и к яркости изображения. Это проиллюстрировано на рис. 1. На рис. 1 представлена три изображения с применением обычной пороговой сегментации, из которой видно, что при низкой яркости фон от объекта практически не отделяется.



Яркость изображения один из фундаментальных параметров прямо определяющий результат сегментация любым из ранее упомянутых методов цветовой сегментации [1, 2]. Если яркость недостаточно высокая, то могут возникать проблемы с определением границ объекта и разделением объектов с похожими цветами или текстурами, аналогично и для избыточно высоких показателях яркости, в этом случае светлые тоны и их границы могут сливаться друг с другом. В этом случае необходимо усложнять алгоритм сегментации изображения и/или выполнять дополнительную предобработку изображения.

Из источников [1, 5] и практики известно, что объекты и границы объектов необходимо отделять при высоком уровне яркости, тем не менее изображения с низким уровнем яркости можно привести к высокому уровню, однако здесь необходимо применять такие методы обработки изображения и параметры этих методов, чтобы сохранить исходную информативность изображения: в виде цвета объекта, границ объектов, размера объекта и т. д.

Условно для компенсации изменений яркости изображения используются два подхода: первый заключается в постоянной ручной настройке пороговой сегментации, второй подход подразумевает разработку динамической сегментации, которая автоматизировано изменяет параметры сегментации, в зависимости от среды и объекта. Для реализации подхода разработан алгоритм, изображенный на рис. 2.

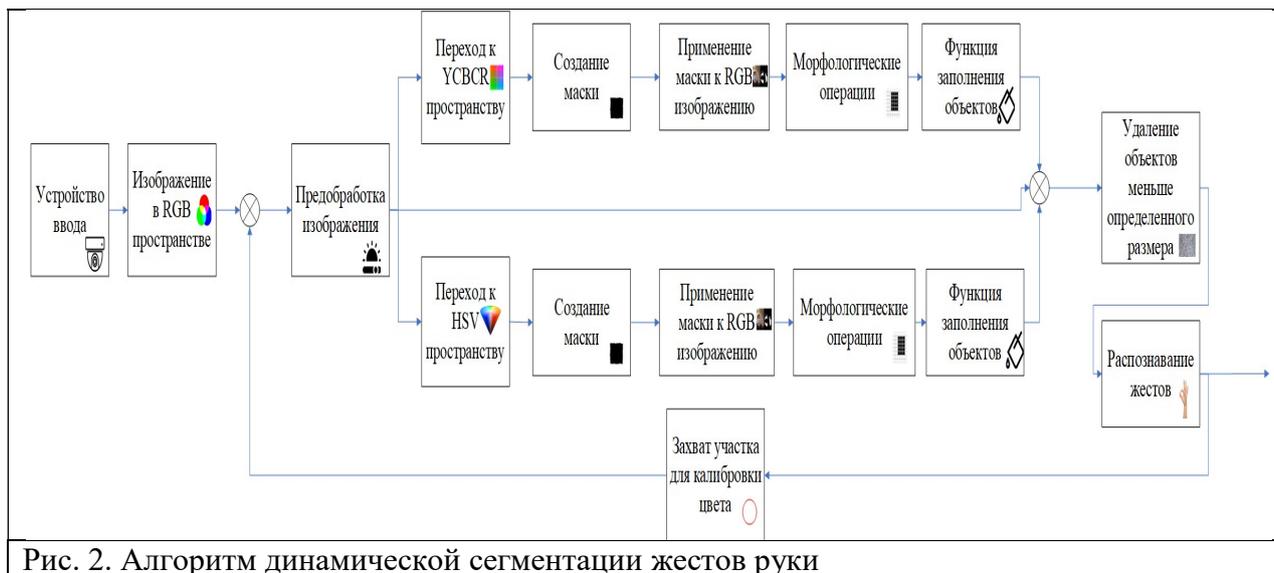


Рис. 2. Алгоритм динамической сегментации жестов руки

Устройства ввода, в качестве которого выступает камера фиксирует видеоизображение, которое представляет из себя набор кадров, изображений в RGB-пространстве, блок предобработки изображения, нормализует уровни яркости на участках изображения, применяет билатеральную фильтрацию и передает полученное изображение в блоки YCBCR и HSV-пространства, после чего создаётся маска и применяется к RGB-изображению, блок морфологических операций, выполняет операции «открытия» и «закрытие», который представляет собой операции дилатации и эрозии в обратных последовательностях. Функция заполнения объектов [6, 8] заполняет оставшиеся «пустоты» в объектах. Выходы этих блоков суммируются тем самым получается одна маска, из которой дополнительно удаляются объекты меньше определенных размеров, после чего маска применяется к RGB-изображению и полученный результат передаётся на блок распознавание жестов [7]. Если необходимо можно провести перекалибровку порогов сегментации, в таком случае человеко-машинный интерфейс предложит пользователю поместить руку на расстоянии 10 см от центра камеры, после чего алгоритм автоматически пересчитает пороги для дальнейшей сегментации жестов руки. Морфологическая операция дилатации выполняется с использованием матрицы предоставленной в (1).

$$dil_m = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \end{pmatrix}. \quad (1)$$

Морфологическая операция эрозии выполняется с использованием матрицы, представленной в (2).

$$er_m = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}. \quad (2)$$

В (1) и (2) единицы — это белые пиксели, а нули — чёрные. Выбор размерности и заполнение матриц проводилось на основании литературы [3, 4], а также эмпирических результатов.

В блоке предобработки изображения, применяются две функции: функция нормализации яркости и функция билатеральной фильтрации. Нормализации яркости необходима для устранения «засветов» и «пересветов» на изображении, которые влияют на результат сегментации изображения, это можно выполнить, используя метод адаптивной эквализации гистограмм, пример изображения до и после представлен на рис. 3.

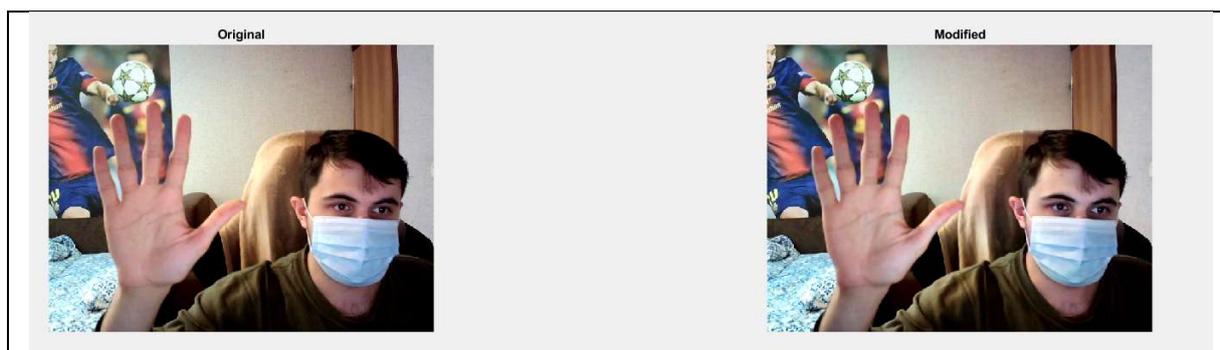


Рис. 3. Пример нормализации яркости изображения, слева исходное изображение, справа модифицированное

Из рис. 3, видно, что из исходного изображения на участках кожи значительно уменьшилась «краснота» кожи, вызванная наличием на потолке помещения красное отражаемой поверхности (натяжной потолок красного цвета). Билатеральная фильтрация по своей сути является модификацией фильтрации Гаусса, однако её главным отличием является, то, что она не размывает границы объектов, тем самым сохраняя контуры объектов.

Функция заполнения объектов заполняет оставшиеся полые участки на всех объектах, найденных на изображении, пример выполнения представлен на рис. 4.



Рис. 4. Пример заполнения объекта, слева исходный объект, справа заполненный

Захват участка для калибровки создаёт маску на участке видеоизображение, после нажатия определенной клавиши или спустя заданное время создается, отбирается четыре последних кадра, затем формируется новый кадр значение каждого пикселя, которого в выделенном участке фильтруется с помощью медианного фильтра из раннее отобранных кадров. На основании полученного кадра формируется пороговые значение для цветových пространств YCBCR и HSV. Это осуществляется за счет преобразование RGB-изображения в YCBCR и HSV-пространство, после чего изображение в каждом из пространств разбивается на три канала, например для HSV, это H, S и V-каналы, для каждого из канала устанавливается свой порог.

### Полученные результаты и их обсуждение

Проведём испытание полученной динамической сегментацией при различных уровнях яркости рис. 5–7, на которых слева вверху представлена оригинальное изображение, сверху справа и снизу слева маски изображения в цветových пространствах HSV и YCBCR соответственно, и снизу справа итоговая маска.

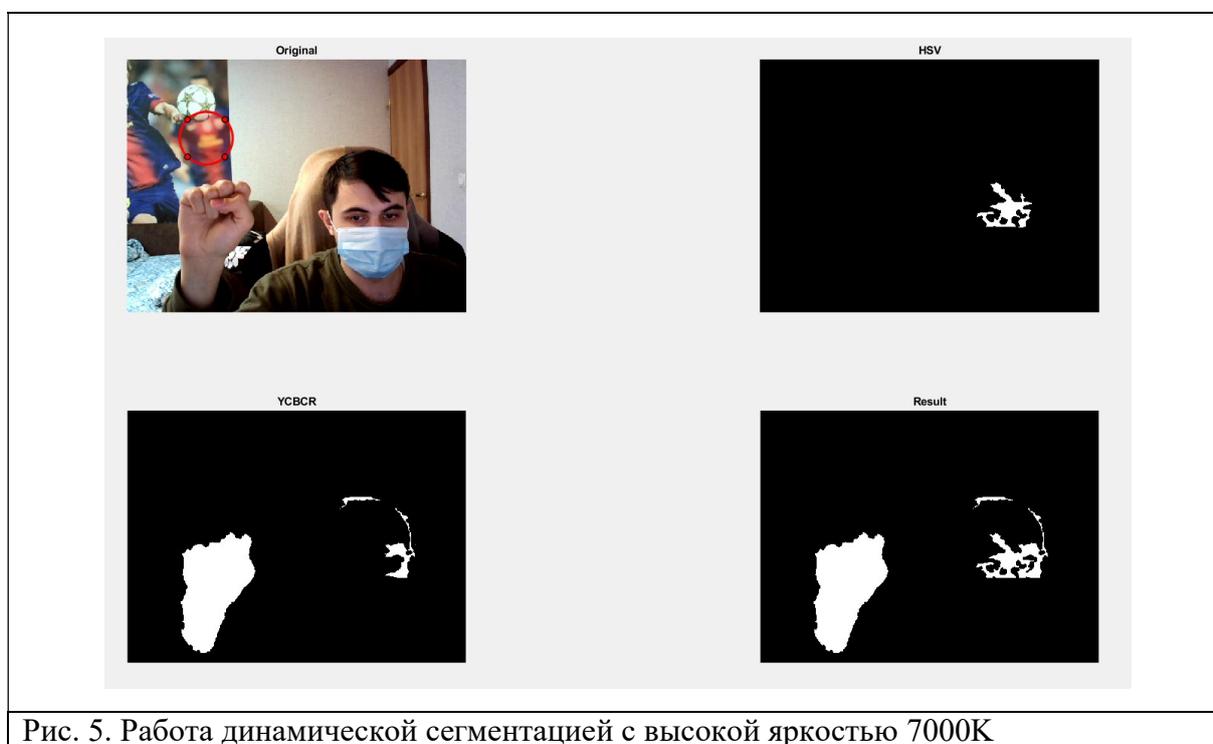


Рис. 5. Работа динамической сегментацией с высокой яркостью 7000К

Из рис. 5, видно что итоговая маска отделяет объект от фона, кроме того условия данного эксперимента наглядно демонстрируют оправданность применения комбинации двух масок, т.к. в этом случае маска HSV не выделяет объект в отличии от маски YCBCR.

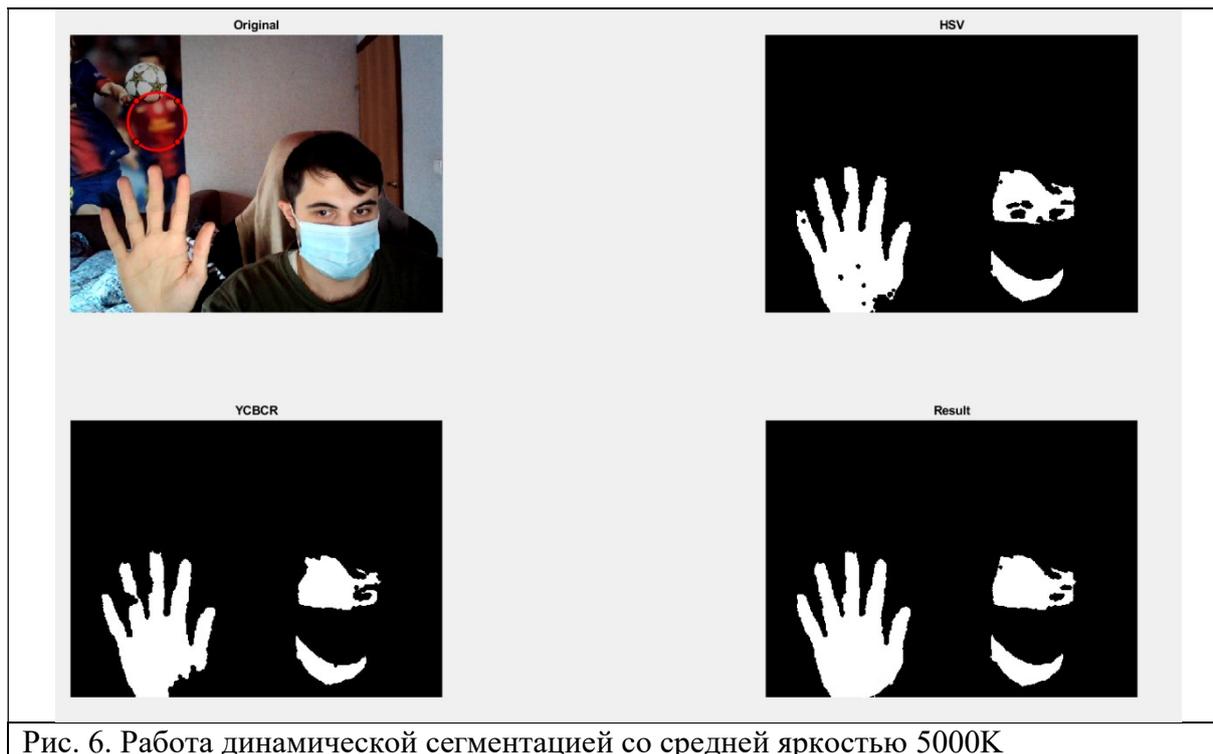


Рис. 6. Работа динамической сегментацией со средней яркостью 5000К

Из рис. 6, видно, что при среднем уровне яркости алгоритм динамической сегментации отделяет фон от объектов белыми пикселями на бинарном изображении.

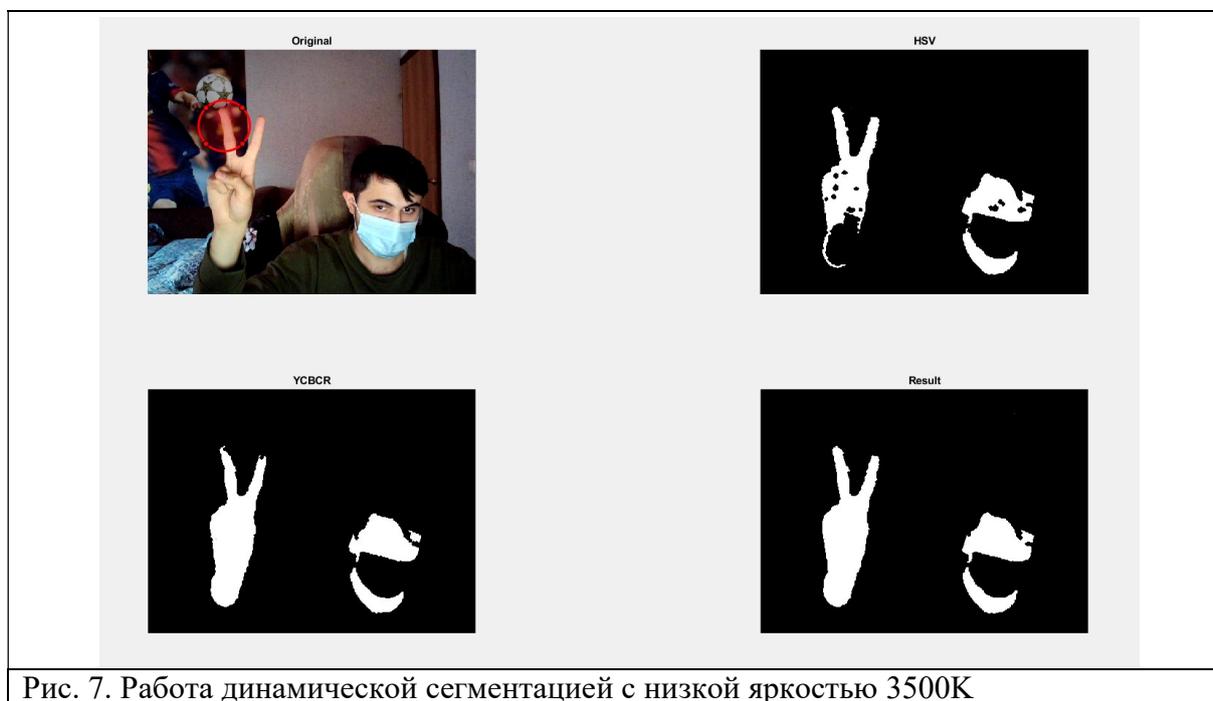


Рис. 7. Работа динамической сегментацией с низкой яркостью 3500К

Из рис. 7 можно сделать вывод, что, как и в случаях с высокой и средней яркостью, объекты отделяются от фона белыми пикселями.

Сравним работу статической и динамической цветовой сегментации при различных уровнях яркости, двух различных комнатах рис. 8.

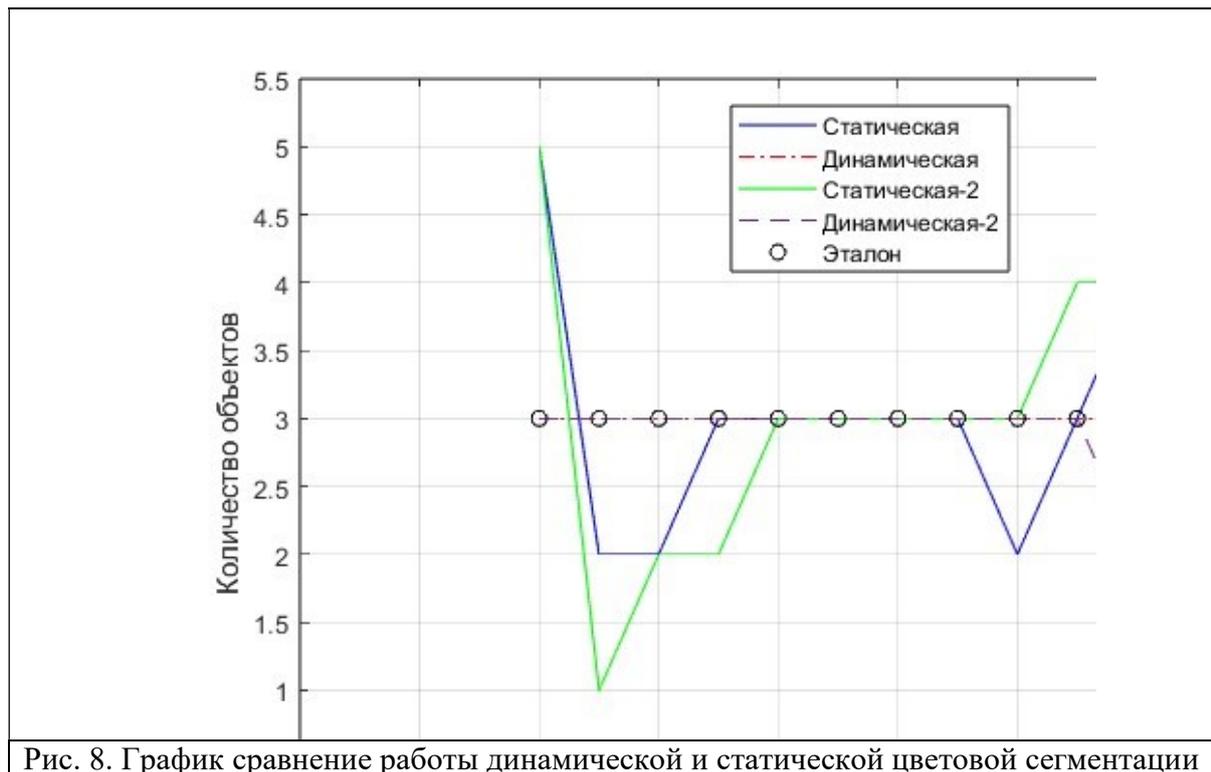


Рис. 8. График сравнение работы динамической и статической цветовой сегментации

Из рис. 8 видно, что при изменении уровня яркости, наиболее близкий к эталону результат показывает динамическая сегментация, отмеченные в легенде «Динамическая» и «Динамическая-2», что соответствует результатам сегментации в двух разных помещениях. Статическая сегментация настраивалась под параметры первой комнаты, динамическая сегментация настраивается независимо от комнаты, поэтому кривая «Статическая-2» имеет больше всего различий с эталоном. Эталон это реально количество объектов на изображении. Отклонение между кривыми «Статическая» и «Статическая-2» связаны с тем, что во втором помещении потолок является отражаемой поверхностью красного цвета, которая дополнительно создаёт помехи в работе алгоритма.

## Вывод

Статическая сегментация не обладает достаточною робастностью по отношению к изменению параметра яркости, в отличие от динамической сегментации, поэтому целесообразно использовать динамическую сегментацию в качестве блока сегментации в системе распознавания жестов руки [6].

Изучение влияния яркости на динамическую сегментацию изображения является важным для множества областей, включая медицину, компьютерное зрение, фотографию, которые используют сегментацию для упрощения и анализа изображений. На основании данного исследования можно сделать вывод, что яркость изображения является важным фактором для динамической сегментации изображения, который следует учитывать в процессе обработки изображений.

При работе с изображениями рекомендуется использовать яркие примеры в диапазоне от 3000 до 6500 К для повышения качества сегментации.

## Конфликт интересов

Авторы статьи заявляют, что у них нет конфликта интересов по материалам данной статьи с третьими лицами на момент подачи статьи в редакцию журнала, и им ничего не известно о возможных конфликтах интересов в настоящем со стороны третьих лиц.

## Список литературы

1. Андреев С. Ю., Небаба С. Г., Макаров М. А. Подготовка изображений лиц в видеопотоке к распознаванию и фильтрация неинформативных изображений // Томский политехнический университет - Томск: 2014.
2. Визильтер Ю. В. Обработка и анализ изображений в задачах машинного зрения / Ю. В. Визильтер и др. - М.: ФИЗМАТКН, 2010. – 672 с.
3. Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. - М.: Техносфера, 2012. - 1104 с.
4. Емельянов С. В. Информационные технологии и вычислительные системы. Вычислительные системы. Компьютерная графика. Распознавание образов. Математическое моделирование. Выпуск №2, 2015 / С. В. Емельянов. - Москва: Мир, 2015. - 662 с.
5. Сингх С., Прибыльский А. В. Оценка эффективности алгоритма сегментации изображения при контролируемом уровне освещения // КомТех2022 / Сборник трудов – Ростов-на-Дону - Таганрог: 2022.
6. Сингх С., Прибыльский А. В. Классификация жестов на основе адаптивной сегментации видеокadra // Информационные технологии, системный анализ и управления / Сборник трудов – Ростов-на-Дону - Таганрог: 2022/
7. Фу К. Структурные методы в распознавании образов. - М.: Мир, 2005. – 144 с.
8. Яне Б. Цифровая обработка изображений / Б. Яне. - М.: Техносфера, 2007. – 584 с/

## Hand gestures recognition algorithm based on method of dynamic color segmentation of video frame

Singh S., Pribylsky A. V.

*Southern Federal University, Institute of Radio Engineering Systems and Control  
347922, Rostov region, Taganrog, Nekrasovsky per. 42*

Presented work looking for the problem of segmenting objects on a video image with using dynamic color segmentation of a video frame for further classification of segmented objects. The dynamic color segmentation method uses algorithms for color calibration, extracting information about the color properties of an object on video, preprocessing the video frame, then creating a morphologically processed mask and applying it to the incoming video stream. The resulting mask is formed on the basis of two masks formed from images in HSV and YCBCR color spaces. The goal is to research the methods of video frame color segmentation and conduct experiments for use in a hand gesture recognition system. During the experiments, an algorithm for dynamic image segmentation was developed, images with segmentation results were obtained using static and dynamic methods of color segmentation of the image, a graph was given for comparing the work of static and dynamic segmentation at different brightness levels, a description of the images and graphs obtained was given, and a conclusion was drawn from the results of the experiment.

*Keywords:* color segmentation, MatLab, video frame, filtration, morphological processing.