

ЦВЕТОВАЯ СЕГМЕНТАЦИЯ ФОТОИЗОБРАЖЕНИЙ, ПОЛУЧЕННЫХ АВТОНОМНЫМ НЕОБИТАЕМЫМ ПОДВОДНЫМ АППАРАТОМ

А.С. Маметьев, А.М. Павин

В статье описывается разработанный авторами алгоритм сегментации цветных фотоизображений, полученных системой технического зрения автономного необитаемого подводного аппарата. Целью сегментации является последующее выделение (детектирование) на изображении интересующего объекта, обладающего специфичной цветовой палитрой в сравнении с другими (ложными) объектами поверхности дна. Суть разработанного алгоритма заключается в построении гистограмм распределения частоты повторения значений пикселей в цветовой модели тон–насыщенность–яркость (*hue–saturation–value*). Соотношение таких гистограмм детектируемого (искомого) объекта и преимущественного фона (ложных объектов) позволяет сегментировать фотоизображение и распознать на нем интересующий объект. После применения процедуры сегментации производится определение координат и направления детектируемого объекта. Для этих целей в работе используется метод линейной аппроксимации, примененный к совокупности пикселей, прошедших через разработанную процедуру фильтрации. Суть данной процедуры заключается в удалении определённого количества весов пикселей по строкам и по столбцам до тех пор, пока с изображения не будет удалена необходимая сумма весов пикселей. Естественным следствием такой процедуры является полная очистка изображения, что говорит об отсутствии искомого объекта на нем. Преимуществами разработанного алгоритма являются высокая скорость обработки фотокадров (требуется всего один проход по пикселям изображения), скорость обучения (достаточно одного цикла по всем пикселям всех изображений из обучающей выборки) и простота реализации. Приведены примеры работы алгоритма на реальных фотоизображениях, полученных фотосистемой малогабаритного подводного аппарата, предназначенного для обучения студентов и участия в спортивных мероприятиях по подводной робототехнике.

Ключевые слова: цветовая сегментация, обработка изображений, распознавание образов, детектирование объектов, бинаризация изображений, системы технического зрения, автономные подводные аппараты

Введение

Несмотря на повсеместное применение искусственных нейронных сетей для решения задачи сегментации фотоизображений существуют области, где уместно применение классических подходов из теории распознавания образов. Примером может служить обработка фотоизображений с использованием слабых вычислительных ресурсов на процессорах с низким потреблением энергии и небольшим выделением тепла. Подобные условия часто возникают при создании мобильных робототехнических комплексов, например малогабаритных автономных необитаемых подводных аппаратов (АНПА) [1, 2].

Разработанный алгоритм сегментации создавался как составная часть системы технического зрения АНПА *Geek* (рис. 1). Сам аппарат является совместной разработкой студентов, аспирантов и специалистов из Института проблем морских технологий (ИПМТ ДВО РАН) [3] и Дальневосточного федерального университета (ДФУ) [4]. Назначение аппарата – обкатка новых и известных решений в области разработки систем АНПА, обучение студентов основам подводной робототехники, а также участие студенческой команды в российских [5, 6] и международных соревнованиях [7–10] в классе АНПА. В рамках студенческих соревнований системой технического зрения АНПА должны решаться задачи обна-

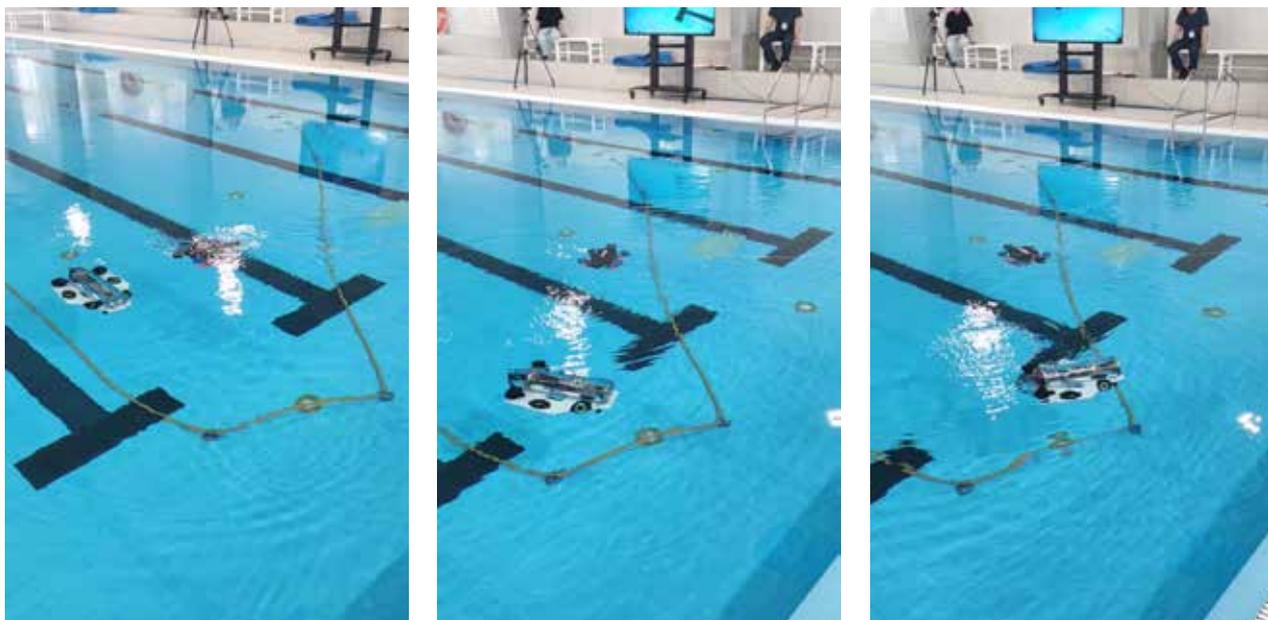


Рис. 1. АНПА Geek во время выполнения задания движения вдоль контрастной линии (шланга)

ружения и локализации как малогабаритных (помещающихся в кадр фотосистемы), так и протяженных объектов различной формы и цветовой палитры.

Перечисленные примеры распознаваемых объектов в большинстве случаев обладают характерной цветовой палитрой (рис. 2). Однако простое пороговое разделение в наиболее популярных цветовых моделях *RGB* (*red, green, blue* – красный, зелёный, синий), *HSV* (*hue, saturation, value* – тон, насыщенность, яркость) или *HSL* (*hue, saturation, lightness* – тон, насыщенность, светлота) часто не позволяет получить приемлемое качество детектирования из-

за изменений условий съемки (освещения, яркости, контрастности и тона как объекта, так и фона).

■ Анализ существующих решений

Метод пороговой бинаризации

Для разделения пикселей изображения на классы «полезные» (точки объекта) и «фоновые» (посторонние) в простейших случаях может использоваться метод пороговой бинаризации (*threshold method*) [11–13]. Данный метод разделяет используемое цве-

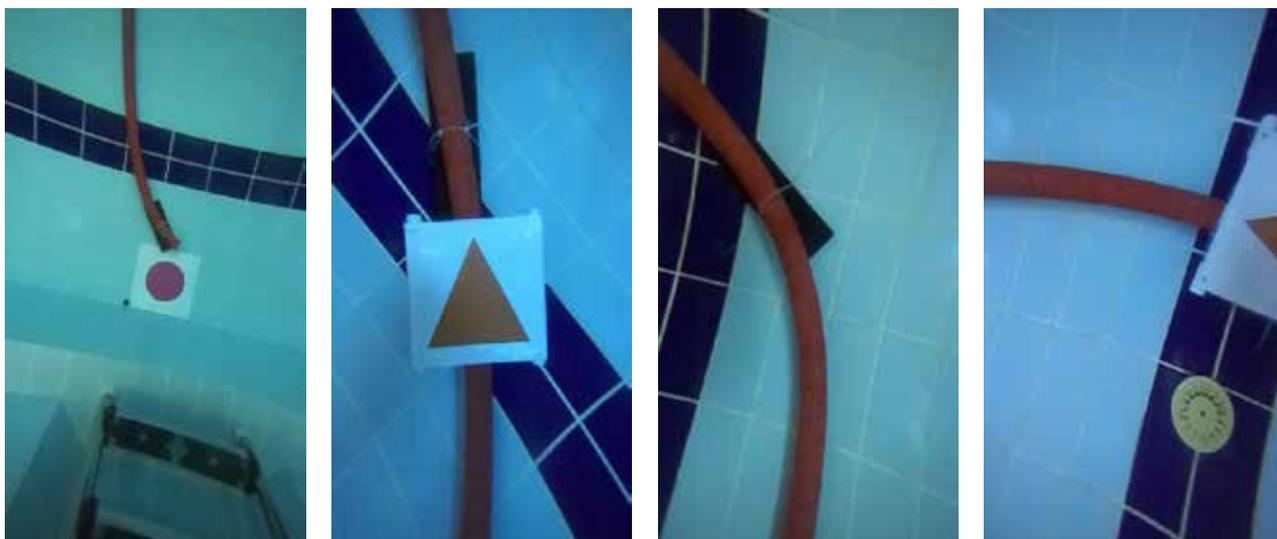


Рис. 2. Пример фотоизображений одного и того же объекта (шланг на дне бассейна) в различных условиях съемки на протяжении одной миссии подводного аппарата

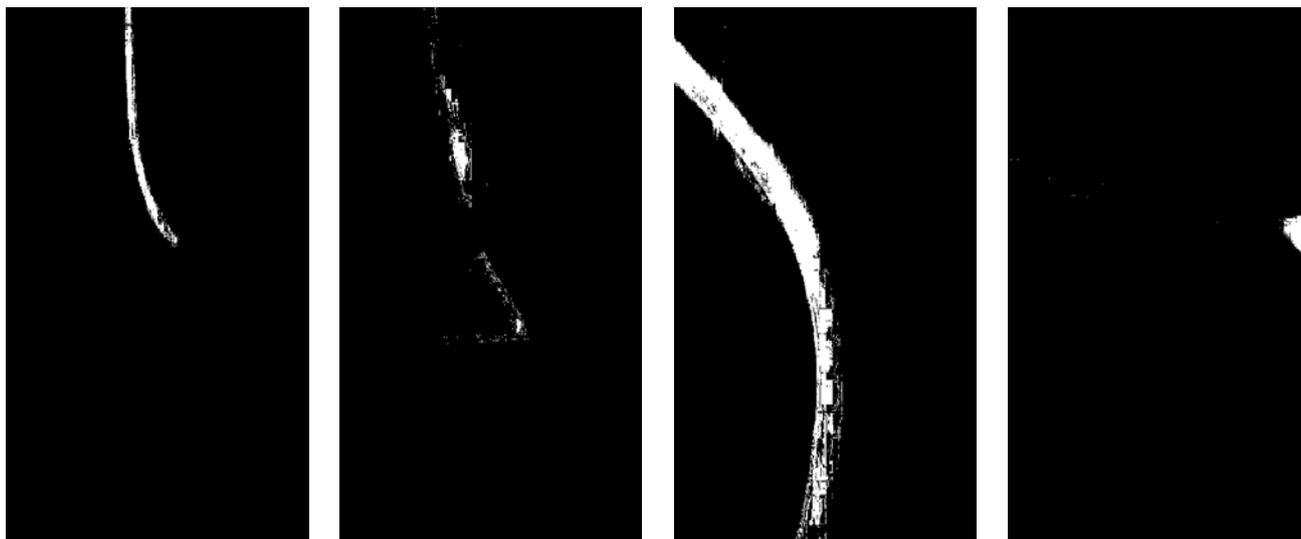


Рис. 3. Пример распознавания объектов методом пороговой бинаризации по параметрам H , S и V

товое пространство на две части заданными параметрами и соотносит каждый пиксель изображения к одной из категорий, формируя тем самым бинарную маску. Плюсом данного метода является крайняя простота реализации и высокая скорость выполнения с асимптотикой $O(N)$. Однако ключевым минусом является нестабильное поведение в условиях переменного освещения, особенно при наличии каких-либо бликов или других дефектов съемки, характерных для режимов функционирования подводного аппарата. Кроме того, данный метод предоставляет только бинарное решение, из-за чего незначительные ложноположительные срабатывания могут серьезно повлиять на предполагаемое местоположение объекта. Примером могут служить фотоизображения на рис. 2, соответствующий результат сегментации которых приведен на рис. 3. На полученных бинаризованных масках можно видеть примеры удачной, удовлетворительной и совсем не удовлетворительной (из-за изменения освещения, глубины и прочих факторов) сегментации красного подводного шланга на фоне дна бассейна и посторонних предметов.

Метод Оцу

Усовершенствованным алгоритмом пороговой бинаризации изображений является метод Оцу (*Otsu method*) [14–16], суть которого заключается в оптимизации представленного выше метода пороговой бинаризации. Для определения оптимального порога в качестве функции ошибки выступает взвешенная сумма дисперсий двух определяемых классов. Минимизировав ошибку, можно получить необходимый порог. В отличие от предыдущего, данный метод

ограничен одним порогом, то есть работает с серошкальными изображениями. Поэтому метод демонстрирует свою работоспособность на изображениях, на которых классы явно отличаются либо по яркости, либо по какому-то другому критерию, например, по значению оператора Собеля [17]. Однако в решаемых АНПА задачах, связанных с обнаружением подводных объектов, классы часто не отличаются друг от друга явным образом по подобным простейшим критериям.

Метод водораздела

Одним из самых популярных алгоритмов сегментации изображений, является метод водораздела (*watershed method*) [18–20]. Его основа состоит в получении серошкального изображения каким-либо алгоритмом (зачастую оператором Собеля) и в последующем нахождении границ между участками с локальными минимумами. Хотя метод позволяет достаточно удачно отделить большое количество различных объектов друг от друга, его существенным недостатком является образование большего количества классов, чем требуется. В контексте задачи нахождения пикселей конкретного объекта это является серьезной проблемой, так как выяснить, какой именно класс необходимо рассматривать, зачастую невозможно.

Использование нейросетей

Другим популярным методом решения задачи сегментации изображения является использование сегментационных нейронных сетей [21–25]. Возможно как использование готовых продуктов, например

YOLO [21], так и обучение собственной нейросети. Данный метод является достаточно эффективным на практике, так как позволяет учитывать, в сущности, любые характерные черты объекта, а не только цвет. В ряде случаев форма объекта может сильно улучшить качество распознавания. Однако подход обладает некоторыми очень серьезными недостатками, значительно ограничивающими его применимость в рассматриваемых задачах. В первую очередь это высокие требования к вычислительным мощностям бортового компьютера и, как правило, необходимость использования нейронных (*NPU*) или графических (*GPU*) процессоров. Используемые в подводных аппаратах процессоры, как правило, маломощные, так как ограничены в тепловыделении и энергозатратах. В результате обработка изображений нейросетью без применения *NPU* или *GPU* порой составляет несколько кадров в секунду, что заметно затрудняет её использование в контуре системы управления движением (работающем на частоте 10–20 Гц) подводного аппарата при решении задач отслеживания (трекинга) протяженных объектов и задач динамической стабилизации над распознанным объектом. В этой связи при практическом применении нейронных сетей для обработки подводных изображений многие исследователи часто заменяют решение задачи семантической сегментации (определение принадлежности) задачей детектирования обрамляющей рамки вокруг искомого объекта [24, 25], что далеко не всегда удобно на практике для определения координат и направления существенно вытянутых (протяженных) объектов. Кроме того, к минусам применения нейросетей следует отнести сложность корректной настройки и обучения их моделей в условиях нехватки времени, а также потребность в большой обучающей выборке, которая во время решения практических задач часто отсутствует или сильно не сбалансирована.

Обнаруженные недостатки исследованных подходов в контексте специфики АНПА и его назначения побудили разработать алгоритм сегментации фотоизображений, сочетающий в себе простоту реализации, высокую скорость обучения и обработки данных, низкие требования к вычислительным ресурсам и достаточно высокие показатели надежности.

■ Анализ гистограмм распределения значений пикселей распознаваемого объекта и фона

Суть разработанного алгоритма заключается в построении гистограмм распределения значений пикселей в цветовой модели *HSV*. Рассмотрим значе-

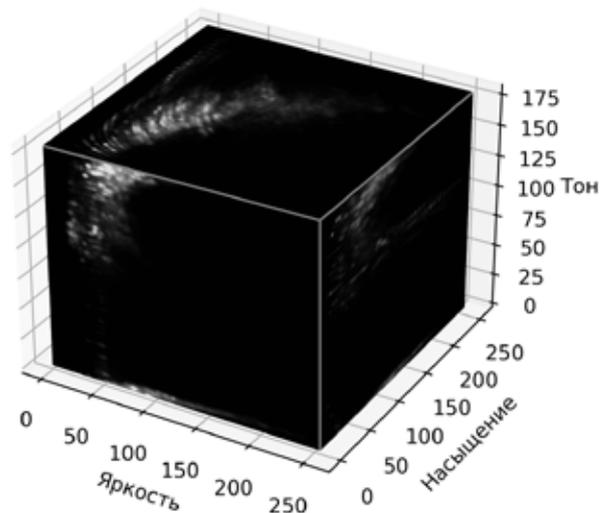


Рис. 4. Проекция трехмерной гистограммы распределения пикселей детектируемого объекта на плоскостях *SV*, *HV* и *HS*

ние гистограмм пикселей, принадлежащих объекту (рис. 4), и пикселей, принадлежащих фону (рис. 5). В качестве примера детектируемого объекта в данной статье будет использоваться подводный протяженный шланг красного цвета, уложенный на дне бассейна. Похожие объекты зачастую используются при проведении российских и международных студенческих соревнований по подводной робототехнике в классе АНПА в задачах, связанных с движением аппарата по визуальным ориентирам. Применяемая для сегментации цветовая модель *HSV* обладает тремя каналами, поэтому сама гистограмма распределения значений пикселей также является трехмерной, которую удобно анализировать по проекциям каналов на плоскости *HS*, *HV* и *SV*.

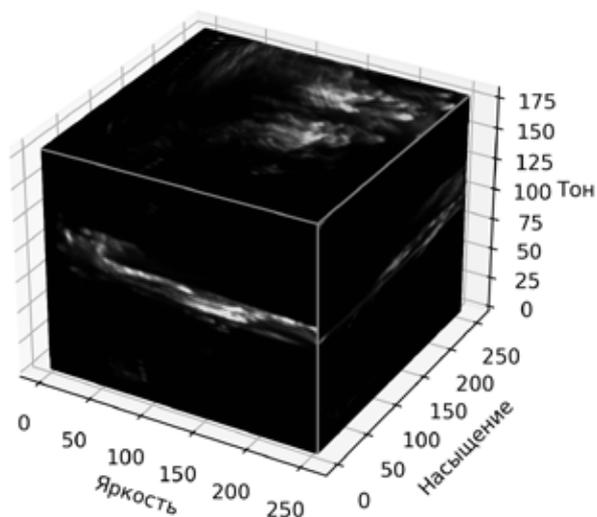


Рис. 5. Проекция трехмерной гистограммы распределения пикселей детектируемого не объекта (фона) на плоскостях *SV*, *HV* и *HS*

На рис. 4, 5 приведены проекции гистограмм детектируемого объекта и фона на плоскостях SV , HV и HS . Гистограммы построены по 28 изображениям (похожим на рис. 2) обучающей выборки, использовавшейся в условиях проведения соревнований по подводной робототехнике. В представленных проекциях каждый пиксель плоскости куба отражает нормированное максимальное количество пикселей, обладающих соответствующей комбинацией значений HS , HV и SV . Можно видеть, что как пикселям объекта, так и пикселям фона соответствуют явно выраженные области в пространстве HSV (светлые области). При этом форма области объекта не является прямоугольной, в связи с чем не может быть описана, например, пороговыми значениями по каждому из каналов в отдельности.

Анализ приведённых на рис. 4, 5 данных показывает, что:

1. Пиксели объекта однозначно локализируются в определённой области пространства HSV , что, согласно гипотезе компактности, позволяет сделать вывод о возможности выделения такого объекта в данном пространстве признаков.

2. Само пространство HSV заполнено точками объекта неравномерно. В частности, на приведённых гистограммах наблюдается множество локальных максимумов и минимумов концентрации точек объекта и фона. Наличие подобных локальных экстремумов объясняется неравномерностью распределения значений точек на реальных фотоизображениях.

Чтобы избавить приведённые выше гистограммы от «мнимых» экстремумов, была уменьшена дискретизация каждой из осей H , S и V до получения приемлемого результата. В представленных данных ось H уменьшалась до 60 дискретов, S – до 16 дискретов, V – до 16 дискретов, что являлось субъективным выбором исследователей, но обеспечивало необходимый результат. Одновременно с этим был ограничен сам максимум на гистограмме (до значения 50% от максимального значения для объекта и 20% для фона), что в конечном итоге позволило получить более равномерные области объекта и фона (рис. 6).

На рис. 6 приведены совмещённые гистограммы объекта и фона, где плотность точек объекта пропорциональна красному каналу, а фона – синему. Можно видеть, что пиксели объекта и фона преимущественно делят различные области пространства HSV , имея в некоторых точках плавные переходы (спорные пиксели). Финальная принадлежность пикселя объекту определяется исходя из соотношений:

$$P_{obj}^{sat} = \begin{cases} P_{obj} / \alpha & P_{obj} \leq \alpha \\ 1 & P_{obj} > \alpha \end{cases} \quad (1)$$

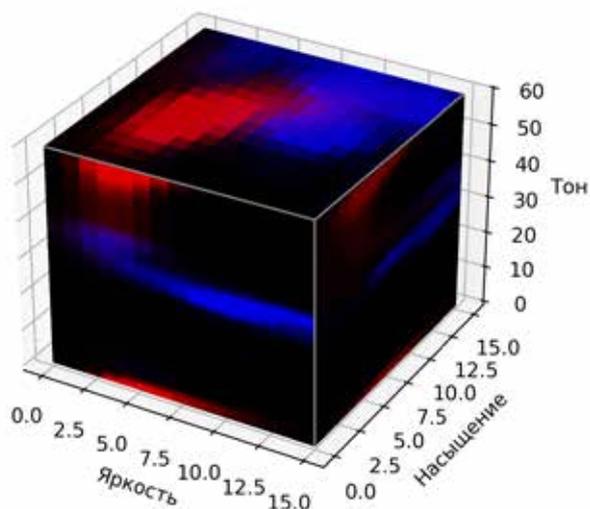


Рис. 6. Проекция гистограммы распределения пикселей детектируемого объекта (красный канал) и фона (синий канал) на плоскостях SV , HV и HS

где P_{obj} – трёхмерная гистограмма объекта для всего цветового пространства HSV , значения $P_{obj}(h,s,v)$ находятся в диапазоне $[0,1]$, α – коэффициент алгоритма, являющийся первичным насыщением гистограммы объекта.

Иными словами, все значения гистограммы, превышающие величину α , будут приравнены к 1; все прочие значения будут нормированы в соответствии с параметром α . На практике используется значение 0,5.

$$P_{bg}^{sat} = \begin{cases} P_{bg} / \alpha & P_{bg} \leq \beta \\ 1 & P_{bg} > \beta \end{cases}, \quad (2)$$

где P_{obj} – трёхмерная гистограмма фона для всего цветового пространства HSV , β – коэффициент алгоритма, являющийся первичным насыщением фона, аналогично α . На практике используется значение 0,2.

$$P_{sub} = P_{obj}^{sat} - P_{bg}^{sat}, \quad (3)$$

$$P = \begin{cases} (P_{sub} - \gamma) / (1 - \gamma) & P_{sub} \geq \gamma \\ 0 & P_{sub} < \gamma \end{cases}, \quad (4)$$

где P – итоговая трёхмерная матрица вероятностей принадлежности элементов цветового пространства HSV объекту, γ – коэффициент алгоритма, являющийся насыщением снизу итоговой вероятности принадлежности пикселя объекту. Все значения, меньшие γ , будут приравнены к 0; остальные значения будут нормированы от 0 до 1. На практике используется значение 0,1.

■ Алгоритм сегментации изображений

В основе разработанного алгоритма лежат вышеприведённые соотношения (1–4). Применение данных соотношений ко всем пикселям входных изображений (рис. 2) даёт следующий результат: рис. 7 – светлые участки соответствуют высокой вероятности принадлежности пикселя выделяемому на изображении объекту, тёмные – низкой, синий цвет – участки с полностью нулевым значением.

Можно видеть, что в большинстве случаев яркие точки действительно соответствуют пикселям объекта, в то время как точки фона обладают либо слабой яркостью (серые и черные области), либо строгим нулем (на изображении обозначен синим цветом). Для определения координат объекта на изображении производился расчёт среднего взвешенного всех точек изображения с весами, равными P (4). В ходе экспериментов выяснилось, что на изображении может присутствовать небольшой процент пикселей, ложно определяемых как «пиксели объекта» (часто обладающих почти нулевыми значениями, рис. 7 – черные и темно-серые области). Чтобы данные точки не могли существенно сместить координаты детектируемого объекта, в алгоритм была добавлена дополнительная процедура фильтрации. Суть процедуры заключается в удалении (занулении) определённого количества весов P по строкам – сверху и снизу, и по столбцам – слева и справа. В ходе фильтрации последовательно выбираются и удаляются строка и столбец с минимальной суммой весов до тех пор, пока с изображения не будет удалена необходимая сумма яркостей удаленных точек. Естественным следствием такой

процедуры является полная очистка (обнуление) весов P в случае слишком малого их суммарного количества, что говорит об отсутствии искомого объекта на изображении. Ещё одним следствием приведенной процедуры является косвенное определение границ (*bounding box*) искомого объекта на изображении.

На рис. 8 можно наблюдать окончательный результат работы алгоритма. Для следования вдоль линии решается задача обнаружения объекта на сегментированном изображении с помощью процедуры линейной аппроксимации. Подобный подход позволяет обнаружить как усредненные координаты искомого объекта, так и его направление. Красная прямая на изображении обозначает полученное линейное приближение точек объекта. При обнаружении малогабаритных объектов на фотоизображениях, для которых требуется определить местоположение, центр может быть определен как среднее взвешенное, а размер – как дисперсия. Светло-голубым цветом на рис. 8 обозначены удалённые в результате фильтрации ложноположительные пиксели. Можно видеть, что для приведенной выборки алгоритм верно выделяет искомым объект с определением его координат и направления.

Разработанный алгоритм был реализован на языке программирования *Python* с использованием библиотеки *Numpy* и внедрён в систему технического зрения АНПА *Geek*. Результаты тестирования алгоритма и апробации в условиях студенческих соревнований показали, что 500 фотоизображений размером 376×762 пикселей на микрокомпьютере *Nvidia Jetson TX2* на одном ядре процессора распознаются примерно за 8 секунд. Таким образом, алгоритм позволяет

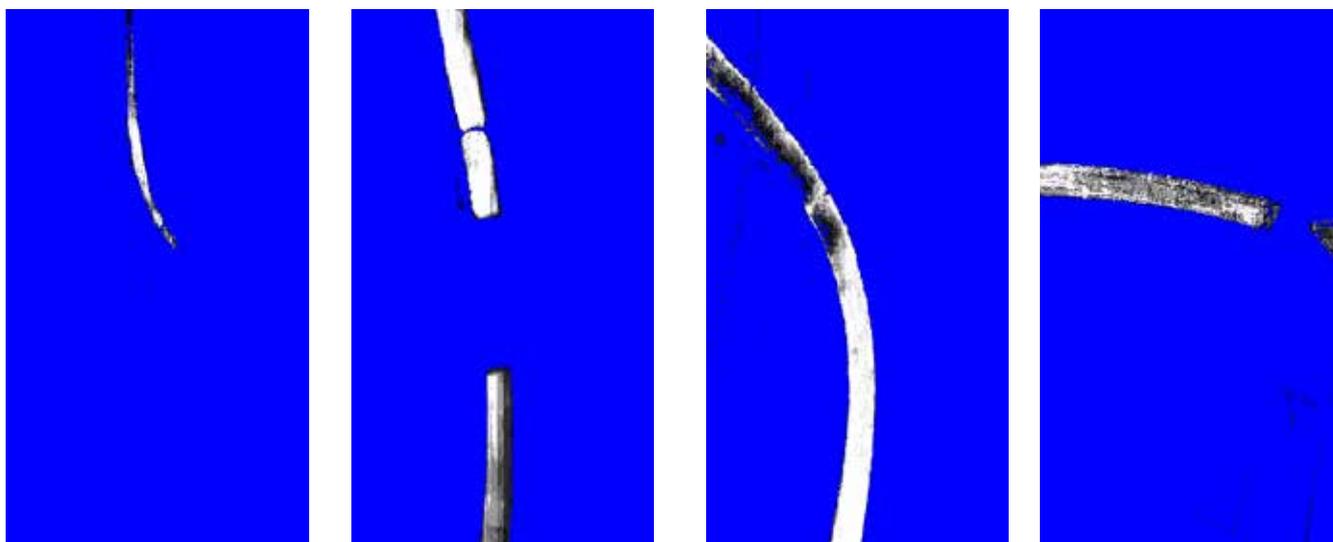


Рис. 7. Результат применения соотношений (1–4) к изображениям, приведенным на рис. 2

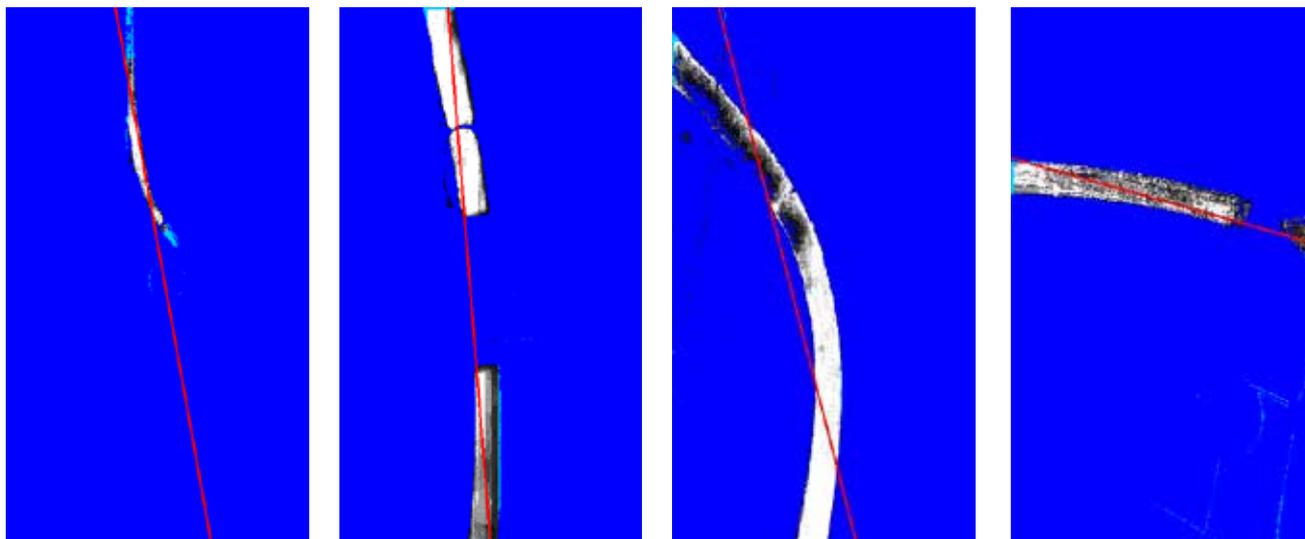


Рис. 8. Итоговый результат детектирования объектов, изображенных на рис. 2

достичь частоты обработки изображений более 60 кадров/секунду на сравнительно маломощном вычислительном процессоре без использования графических и нейросетевых модулей.

В текущей реализации алгоритма не задействовались ядра *GPU* (*graphics processing unit*), *NPU* (*graphics processing unit*) или *CUDA* (*Compute Unified Device Architecture*), применение которых может еще ускорить работу алгоритма на современных вычислителях. Кроме того, в реализации алгоритма использовалась заведомо избыточная точность хранения значений гистограмм – 32 бита, а также исследовались режимы работы с точностью в 8, 16 и 64 бит. Было установлено, что понижение точности до 8 или 16 бит не приводит к хоть сколько-нибудь заметному изменению качества и ускоряет расчёты на архитектуре *x86_64*, но негативно сказывается на производительности для архитектуры *ARM64*, используемой в вычислителях *Nvidia Jetson TX2*. Точность в 64 бита на обеих архитектурах показала небольшое замедление, но также не привела к изменению качества. Вероятность ложноположительных и ложноотрицательных срабатываний на текущий момент досконально не исследовалась и сильно зависит от условий эксплуатации АНПА.

Заключение

В рамках проведённых исследований был разработан алгоритм сегментации цветных фотоизображений, полученных фотосистемой автономного необитаемого подводного аппарата *Geek*. Основу предложенного подхода составляет построение гистограммы распределения пикселей изображения в

цветовом пространстве *HSV* с дальнейшей сегментацией исходного изображения на основе плотности распределения пикселей в построенной гистограмме. К преимуществам алгоритма следует отнести простоту его реализации, скорость обработки данных и возможность быстрой перенастройки (обучения и дообучения) цветовой модели (полученной гистограммы) под новые распознаваемые объекты или под изменившиеся условия съемки. Данные преимущества часто являются определяющими в реальных ситуациях, связанных с разработкой, настройкой и эксплуатацией АНПА. Примером служат соревнования по подводной робототехнике, в которых возможность быстрой перенастройки простого программного обеспечения под изменившиеся условия является ключом к успеху. С применением указанного алгоритма в составе системы технического зрения АНПА *Geek* команда студентов ДВФУ под руководством наставников из ИПМТ ДВО РАН не раз завоёвывала золото на российских [5, 6] и зарубежных [7] соревнованиях. Применение предложенного метода в случае насыщенного неоднородностями фона (часто встречающегося в морских условиях) ограничено в том случае, если детектируемый объект и фон обладают схожими распределениями одновременно по трем каналам *HSV*. К недостаткам разработанного алгоритма также можно отнести тенденцию к ухудшению качества сегментации при тусклом свете. Данный факт обусловлен тем, что в цветовом пространстве *HSV* расстояние между двумя различными цветами уменьшается при уменьшении яркости. Дальнейшее направление работ в данной тематике видится в дополнении предложенного подхода методами, включающими анализ формы объектов на изображении.

СПИСОК ИСТОЧНИКОВ

1. Инзарцев А.В., Киселев Л.В., Костенко В.В., Матвиенко Ю.В., Павин А.М., Щербатюк А.Ф. Подводные робототехнические комплексы: системы, технологии, применение. Владивосток, 2018. 368 с. ISBN 978-5-7311-0486-79.
2. Инзарцев А.В., Павин А.М., Лебедко О.А., Панин М.А. Распознавание и обследование малоразмерных подводных объектов с помощью автономных необитаемых подводных аппаратов // Подводные исследования и робототехника. 2016. №. 2(22). С. 36–43.
3. ИПМТ ДВО РАН. URL: <http://imtp.febras.ru/> (дата обращения: 13.12.2024)
4. FEFU. URL: <https://www.dvfu.ru/en/> (дата обращения: 13.12.2024)
5. Итоги открытого первенства Санкт-Петербурга по морской робототехнике // Новости СПбГМТУ. 2024. URL: <https://www.smtu.ru/ru/viewnews/753/> (дата обращения: 13.12.2024)
6. Аквароботех-2023 // Новости МГУ. 2023. URL: <https://msun.ru/ru/news/id-8897> (дата обращения: 13.12.2024)
7. WUURC 2023. 2023. URL: <https://m.w2urc.org/en/> (дата обращения: 15.08.2023)
8. The results of the MURC-WUURC 2024. 2024. URL: <https://murc.pro/news/469> (дата обращения: 13.12.2024)
9. SAUVC. URL: <https://sauvc.org/> (дата обращения: 13.12.2024)
10. Robosub. URL: <https://robosub.org/> (дата обращения: 13.12.2024)
11. Aqthobilrobbany A., Handayani A.N., Lestari D., Muladi, Asmara R.A., Fukuda O. HSV Based Robot Boat Navigation System // International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM). Surabaya, Indonesia, 2020. 5 p.
12. Niu Z., Li, H. Research and analysis of threshold segmentation algorithms in image processing // Journal of Physics: Conference Series. 2019. Vol. 1237.
13. Kaur G., Singh K. A Comparative Study of Image Segmentation using Thresholding Techniques // Conference: National Conference on Advanced Computing Technologies. Rohtak, India, 2013.
14. Otsu N. A Threshold Selection Method from Gray-Level Histograms // IEEE Transactions on Systems, Man, and Cybernetics. 1979. Vol. 9, No. 1. P. 62–66.
15. Goh T.Y., Basah S.N., Yazid H., Safar M.J.A., Saad F.S.A. Performance analysis of image thresholding: Otsu technique // Measurement. 2018. Vol. 114. P. 298–307.
16. Ershov E.I., Korchagin S.A., Kokhan V.V., Bezmaternykh P.V. A generalization of Otsu method for linear separation of two unbalanced classes in document image binarization // КО. 2021. No. 1.
17. Sobel I., Feldman G. An Isotropic 3x3 Image Gradient Operator. 2014.
18. Beucher S., Lantuéj C. Use of Watersheds in Contour Detection // International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation. 1979.
19. Roerdink J., Meijster A. The Watershed Transform: Definitions, Algorithms and Parallelization Strategies // FUNDINF: Fundamenta Informatica. 2000. Vol. 41.
20. Beucher S. The Watershed Transformation Applied To Image Segmentation // Scanning Microsc. 2000. Vol. 6
21. Mohanapriya S., Mohana Saranya S., Kumaravel T., Sumithra P. Image detection and segmentation using YOLO v5 for surveillance // ACE. 2023. Vol. 8. P. 142–147.
22. Павин А.М., Шилин К.Д. Сегментационная нейронная сеть для обнаружения трубопроводов на эхограммах гидролокатора бокового обзора // Морские интеллектуальные технологии. 2024. №. 4. Ч 1. С. 277–284. DOI: 10.37220/MIT.2024.66.4.033.
23. Ansari M., Bhosale S., Choudhary A. Semantic Segmentation using Convolutional Neural Networks // International Journal of Engineering Research in Computer Science and Engineering. 2023. Vol. 10. P. 31–34.
24. Li Y, Zhang X, Shen Z. YOLO-Submarine Cable: An Improved YOLO-V3 Network for Object Detection on Submarine Cable Images // Journal of Marine Science and Engineering. 2022. Vol. 10(8). P. 1143. URL: <https://doi.org/10.3390/jmse10081143>.
25. Fikri Arif Wicaksana, Eueung Mulyana, Syarif Hidayat, Rahadian Yusuf Design and Implementation Submarine Cable Object Detection YOLOv4 based with Graphical User Interface (GUI) for Remotely Operated Vehicle(ROV) // International Journal of Advanced Computer Science and Applications (IJACSA). 2023. Vol. 14, No. 9. P. 966–981. URL: https://thesai.org/Downloads/Volume14No9/Paper_101-Design_and_Implementation_Submarine_Cable_Object_Detection.pdf

Справка об авторах

МАМЕТЬЕВ Артемий Сергеевич, техник лаб. Систем технического зрения

Федеральное государственное бюджетное учреждение науки Институт проблем морских технологий им. академика М.Д. Агеева Дальневосточного отделения Российской академии наук

Адрес: Россия, 690091, г.Владивосток, ул.Суханова, д.5а

Область научных интересов: программирование, техническое зрение, подводная робототехника

E-mail: agent002004@gmail.com

ORCID: 0009-0008-4641-4029

ПАВИН Александр Михайлович, к.т.н., в.н.с. лаб. Систем технического зрения

Федеральное государственное бюджетное учреждение науки Институт проблем морских технологий им. академика М.Д. Агеева Дальневосточного отделения Российской академии наук

Адрес: Россия, 690091, г.Владивосток, ул.Суханова, д.5а

Область научных интересов: системы технического зрения, распознавание образов, управление подвижными объектами, НПА, АНПА, ТНПА

E-mail: pavin@bk.ru

ORCID: 0000-0001-8878-7888

WoS ResearcherID: AAO-2252-2020

Scopus AuthorID: 24765383300

SPIN-код: 8565-7211

РИНЦ AuthorID: 508298

COLOR IMAGE SEGMENTATION ALGORITHM FOR AUTONOMOUS UNDERWATER VEHICLES

A.S. Mametev, A.M. Pavin

This paper describes an algorithm for color segmentation of images obtained by the computer vision system of an autonomous underwater vehicle. The main goal of image segmentation is further highlighting (detection) of the desired object, having different color palette in contrast with other (false) objects of the water bottom, in a photo. The essence of the algorithm lies in the construction of histograms of distribution of the frequency of occurrences of pixel values in the hue-saturation-value color model. The ratio of such histograms between the detected object and the predominant background allows to highlight (locate) the target object. After the segmentation process the object coordinates and rotation are calculated. For these purpose linear approximation is used on pixels which were processed with filtration algorithm. The filtration consists of sequential removal of rows and columns with minimal sum from the image weight mask, until the required total sum of weights is deleted. This naturally leads to the full zeroing of the images with low total weight sum, which indicates that the object is not present on this image. The advantages of the developed algorithm include high speed of image processing (only a single pass through the image pixels is required), fast training (single pass through the pixels of the training set images), and ease of implementation. The paper provides examples of the algorithm's performance on real photographic images obtained by an underwater vehicle's imaging system, designed for student training and for participating in underwater robotics sports events.

Keywords: color segmentation, image processing, object detection, photo binarization, computer vision system, underwater robot, AUV.

References

- Inzarcev A.V., Kiselev L.V., Kostenko V.V., Matvienko Ju.V., Pavin A.M., Shherbatjuk A.F. *Podvodnye robototekhnicheskie komplekсы: sistemy, tehnologii, primenenie*. Vladivostok, 2018. 368 s. ISBN 978-5-7311-0486-79. [In Russ.]
- Inzarcev A.V., Pavin A.M., Lebedko O.A., Panin M.A. *Raspoznavanie i obsledovanie malorazmernih podvodnyh ob#ektov s pomoshh'ju avtonomnyh neobitaemyh podvodnyh apparatov // Podvodnye issledovaniya i robototekhnika*. 2016. No. 2(22). S. 36–43. [In Russ.]
- IPMT DVO RAN. URL: <http://imtp.febras.ru/> (data obrashheniya: 13.12.2024) [In Russ.]
- FEFU. URL: <https://www.dvfu.ru/en/> (data obrashheniya: 13.12.2024) [In Russ.]
- Itogi otkrytogo pervenstva Sankt-Peterburga po morskoy robototekhnike // *Novosti SPbGMTU*. 2024. URL: <https://www.smtu.ru/ru/viewnews/753/> (data obrashheniya: 13.12.2024) [In Russ.]
- Akvaroboteh-2023 // *Novosti MGU*. 2023. URL: <https://msun.ru/news/id-8897> (data obrashheniya: 13.12.2024) [In Russ.]
- WUURC 2023. URL: <https://m.w2urc.org/en/> (Access date: 15.08.2023)
- The results of the MURC-WUURC 2024. 2024. URL: <https://murc.pro/news/469> (Access date: 13.12.2024)
- SAUVC. URL: <https://sauvc.org/> (Access date: 13.12.2024)
- Robosub. URL: <https://robosub.org/> (Access date: 13.12.2024)
- Aqthobilrobbany A., Handayani A.N., Lestari D., Muladi, Asmara R.A., Fukuda O. *HSV Based Robot Boat Navigation System // International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*. Surabaya, Indonesia, 2020. 5 p.
- Niu Z., Li, H. *Research and analysis of threshold segmentation algorithms in image processing // Journal of Physics: Conference Series*. 2019. Vol. 1237.
- Kaur G., Singh K. *A Comparative Study of Image Segmentation using Thresholding Techniques // Conference: National Conference on Advanced Computing Technologies*. Rohtak, India, 2013.
- Otsu N. *A Threshold Selection Method from Gray-Level Histograms // IEEE Transactions on Systems, Man, and Cybernetics*. 1979. Vol. 9, No. 1. P. 62–66.
- Goh T.Y., Basah S.N., Yazid H., Safar M.J.A., Saad F.S.A. *Performance analysis of image thresholding: Otsu technique // Measurement*. 2018. Vol. 114. P. 298–307.
- Ershov E.I., Korchagin S.A., Kokhan V.V., Bezmaternykh P.V. *A generalization of Otsu method for linear separation of two unbalanced classes in document image binarization // KO*. 2021. No. 1.
- Sobel I., Feldman G. *An Isotropic 3x3 Image Gradient Operator*. 2014.
- Beucher S., Lantuéj C. *Use of Watersheds in Contour Detection // International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation*. 1979.
- Roerdink J., Meijster A. *The Watershed Transform: Definitions, Algorithms and Parallelization Strategies // FUNDINF: Fundamenta Informatica*. 2000. Vol. 41.
- Beucher S. *The Watershed Transformation Applied To Image Segmentation // Scanning Microsc*. 2000. Vol. 6
- Mohanapriya S., Mohana Saranya S., Kumaravel T., Sumithra P. *Image detection and segmentation using YOLO v5 for surveillance // ACE*. 2023. Vol. 8. P. 142–147.
- Pavin A.M., Shilin K.D. *Segmentacionnaja neironnaja set' dlja obnaruzheniya truboprovodov na jehogrammah gidrolokatora bokovogo obzora // Morskie intellektual'nye tehnologii*. 2024. No. 4. Ch 1. S. 277—284. DOI: 10.37220/MIT.2024.66.4.033. [In Russ.]

23. Ansari M., Bhosale S., Choudhary A. Semantic Segmentation using Convolutional Neural Networks // International Journal of Engineering Research in Computer Science and Engineering. 2023. Vol. 10. P. 31–34.

24. Li Y, Zhang X, Shen Z. YOLO-Submarine Cable: An Improved YOLO-V3 Network for Object Detection on Submarine Cable Images // Journal of Marine Science and Engineering. 2022. Vol. 10(8). P. 1143. URL: <https://doi.org/10.3390/jmse10081143>.

25. Fikri Arif Wicaksana, Eueung Mulyana, Syarif Hidayat, Rahadian Yusuf Design and Implementation Submarine Cable Object Detection YOLOv4 based with Graphical User Interface (GUI) for Remotely Operated Vehicle (ROV) // International Journal of Advanced Computer Science and Applications (IJACSA). 2023. Vol. 14, No. 9. P. 966–981. URL: https://thesai.org/Downloads/Volume14No9/Paper_101-Design_and_Implementation_Submarine_Cable_Object_Detection.pdf

Information about the authors

MAMETEV Artemiy S., Technician of the Computer vision Lab. Institute of Marine Technology Problems, Far Eastern Branch of the Russian Academy of Sciences.

Work address: 5-a, Sukhanov street, Vladivostok, 690950, Russia
Research Interests: programming, computer vision, underwater robotics

E-mail: agent002004@gmail.com

ORCID: 0009-0008-4641-4029

PAVIN Alexander M., Ph.D., Leading researcher, head of the Computer vision Lab.

Institute of Marine Technology Problems, Far Eastern Branch of the Russian Academy of Sciences.

Work address: 5-a, Sukhanov street, Vladivostok, 690950, Russia
Research Interests: computer vision systems, pattern recognition, control systems, UUV, AUV, ROV

E-mail: pavin@bk.ru

ORCID: 0000-0001-8878-7888

WoS ResearcherID: AAO-2252-2020

Scopus AuthorID: 24765383300

SPIN-code: 8565-7211



Recommended citation:

Mametev A.S., Pavin A.M. COLOR IMAGE SEGMENTATION ALGORITHM FOR AUTONOMOUS UNDERWATER VEHICLES. Underwater investigations and robotics. 2025. No. 1 (51). P. 41–50. DOI: 10.37102/1992-4429_2025_51_01_04. EDN: BCTUTW.